

GenCore version 4.5
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM nucleic - nucleic search, using sw model
Run on: August 13, 1999, 21:54:13 ; Search time 909.15 Seconds
(without alignments)
12229.403 Million cell updates/sec

Title: US-09-042-460-1
Sequence: 1 GAATTCGGGTGGAGGCC.....CCGAGCTCGGTACCAAGCTT 3496

Scoring table: IDENTITY_NUC
Searched: 679419 seqs, 1590154680 residues

Database: GenEmbl.*

- 1: gb_bal.*
- 2: gb_ba2.*
- 3: gb_om.*
- 4: gb_ov.*
- 5: gb_pat.*
- 6: gb_ph.*
- 7: gb_pl1.*
- 8: gb_pl2.*
- 9: gb_pr1.*
- 10: gb_pr2.*
- 11: gb_pr3.*
- 12: gb_ro.*
- 13: gb_st.*
- 14: gb_sy.*
- 15: gb_un.*
- 16: gb_vl.*
- 17: gb_wg.*
- 18: gb_yl.*
- 19: gb_zl.*

Pred. No. is the number of
score greater than or equal to
and is derived by analysis of the

Result No. 1
Score Match Length DB ID
3426 98.0 3426 12 AF051911

SUMMARY chance to have a
result being printed,
distribution.

Description

Mus muscu

us-09-042-460-1.rge

KAU9MAL
09/04/1999
SEQID: 12

2	3367.4	96.3	3369	12	AF073311	AF073311 Mus muscu
3	1603.6	45.9	8960	15	AF043739	AF043739 Synthetic
4	1600.8	45.8	4015	11	AF015950	AF015950 Homo sapi
5	1600.8	45.8	4027	11	AF018167	AF018167 Homo sapi
6	1462.6	41.8	7498	12	AF121949	AF121949 Mus muscu
7	468.8	13.4	12213	42	AF114847	AF114847 Homo sapi
8	468.8	13.4	15332	42	AF121948	AF121948 Homo sapi
9	468.8	13.4	26414	42	HSPERT1	AF128893 Homo sapi
10	437.4	12.5	5491	9	AB016767	AB016767 Homo sapi
11	194.6	5.6	201	12	AF029235	AF029235 Homo sapi
12	154.8	4.4	4356	11	AF097365	AF097365 Mus muscu
13	154.8	4.4	2043	11	AF098956	AF098956 Homo sapi
14	122.2	3.5	25138	42	HSPERT2	AF128894 Homo sapi
15	63.6	1.8	3418	10	HSP58KRC	AJ000542 Homo sapi
16	63.6	1.8	1102	8	AF137070	AF137070 Sphenosty
17	61.6	1.8	758	36	AAU60877	U60877 Anopheles a
18	61.2	1.8	339	2	AF114621	AF114621 Unculture
19	59.6	1.7	779	36	AAU60880	U60880 Anopheles a
20	59.4	1.7	779	36	AAU60880	U60880 Anopheles a
21	58.8	1.7	461	2	AF033115	AF033115 Mus muscu
22	58.6	1.7	1242	36	SSU52949	AF018036 Unculture
23	57.8	1.7	1631	2	AF018036	AF018036 Spisula sol
24	57.8	1.7	689	2	AF018036	AF018036 Unculture
25	57.8	1.7	719	2	AF018052	AF018052 Unculture
26	57.8	1.7	719	2	AF018052	AF018052 Unculture
27	57.8	1.7	1199	17	DVU93295	AF018054 Unculture
28	57.8	1.7	1197	17	DVU93295	U93295 Dengue viru
29	57.8	1.7	1200	17	DVU93302	U93302 Dengue viru
30	57.8	1.7	775	36	AAU60876	U60876 Anopheles a
31	57.6	1.6	226	12	MUSFRE	L39119 Mus muscu
32	57.4	1.6	448	36	SRMETPANS	AF001134 Strongylo
33	57.2	1.6	731	2	AF018053	AF018053 Unculture
34	57.2	1.6	1088	2	AF056365	AF056365 Unculture
35	57.2	1.6	230	10	HUMX0133D	L33810 Human (clon
36	57.2	1.6	3942	15	AF073995	AF073995 Vibrio ch
37	56.8	1.6	1528	36	BLA35433	AJ005433 Branchios
38	56.8	1.6	3497	10	BSA35433	D88984 Mus muscu
39	56.8	1.6	1673	12	D88984168	D4443 Mouse mRNa
40	56.8	1.6	573	12	MUSDIP	L22944 Human immun
41	56.8	1.6	3110	17	HVISE364A	AJ010397 Hevea bra
42	55.8	1.6	1229	7	HMU010397	U89924 Mus muscu
43	55.4	1.6	1169	12	ECY10545	Y10545 Cloning vec
44	55	1.6	5025	1	AF056364	AF056364 Unculture
45	55	1.6	1081	2	AF056364	

ALIGNMENTS

RESULT	1	AF051911	Mus musculus telomerase reverse transcriptase mRNA, complete cds.	3426 bp	ROD	02-APR-1998
LOCUS		AF051911				
DEFINITION		AF051911				
ACCESSION		AF051911				
NID		92005591				
VERSION		AF051911.1	GI:3005591			
KEYWORDS						
SOURCE			House mouse.			
ORGANISM			Mus musculus			
REFERENCE			Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia;			
AUTHORS			Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
TITLE			1 (bases 1 to 3426)			
JOURNAL			Greenberg, R.A., Allsopp, R.C., Chin, L., Morin, G.B. and DePinho, R.A.			
REFERENCE			Expression of mouse telomerase reverse transcriptase during			
AUTHORS			Oncogene (1998) In press			
TITLE			2 (bases 1 to 3426)			
JOURNAL			Greenberg, R.A., Allsopp, R.C., Chin, L., Morin, G.B. and DePinho, R.A.			
REFERENCE			Submitted (02-MAR-1998) Microbiology and Immunology, Albert			
TITLE			Einstein College of Medicine, 1300 Morris Park Ave, Bronx, NY			
JOURNAL			10461, USA			
FEATURES			Location/Qualifiers			

Db	1561	GCAGCAGCCCGGGGAAGACCGGTGTCCCGCTGCAGAGCACCGTCTGAGGGAGAGATCC	1620
Qy	1630	TGGCTACGTTCCCTGTCTTGGCGTGATGGACACATACGTTGGTACAGCTGCTTAGGTCATTCT	1689
Db	1621	TGGCTACGTTCCCTGTCTTGGCGTGATGGACACATACGTTGGTACAGCTGCTTAGGTCATTCT	1680
Qy	1690	TTTACATCACAGAGAGCACATTCAGAAAGAACAGCGCTTCTTCTACCGCTAAAGAGTGCTGT	1749
Db	1681	TTTACATCACAGAGAGCACATTCAGAAAGAACAGCGCTTCTTCTACCGTAAAGAGTGCTGT	1740
Qy	1750	GGACCAAGCTGCACAGCATTTGGAGTCAAGGCACACACTTGAGAGAGTCCGGCTACGGGAGC	1809
Db	1741	GGACCAAGCTGCACAGCATTTGGAGTCAAGGCACACACTTGAGAGAGTCCGGCTACGGGAGC	1800
Qy	1810	TGTCACAAGAGAGGTTCAGGCATCACAGAGACACTTGCTAGCCATGCCCATTCTCGAGAC	1869
Db	1801	TGTCACAAGAGAGGTTCAGGCATCACAGAGACACTTGCTAGCCATGCCCATTCTCGAGAC	1860
Qy	1870	TGGCGTTTCATCCCAAGCCCAAGCGCTCGGGCCCAATTGTGAACATAGATTTAGCATGG	1929
Db	1861	TGGCGTTTCATCCCAAGCCCAAGCGCTCGGGCCCAATTGTGAACATAGATTTAGCATGG	1920
Qy	1930	GTACCAAGAGCTTTGGCGAAGGAAGCAGGCCAGCATTTACCCACGCTCTCAAGACTC	1989
Db	1921	GTACCAAGAGCTTTGGCGAAGGAAGCAGGCCAGCATTTACCCACGCTCTCAAGACTC	1980
Qy	1990	TCCTTCAGCATGCTCAACTATGAGCGGACAAAACATCCTCACCTTATGGGCTCTCTGTAC	2049
Db	1981	TCCTTCAGCATGCTCAACTATGAGCGGACAAAACATCCTCACCTTATGGGCTCTCTGTAC	2040
Qy	2050	TGGGTATGAATGACATCTACAGGACCTGCGGGGCCCTTTGTGCTGGGTGCGTGTCTGG	2109
Db	2041	TGGGTATGAATGACATCTACAGGACCTGCGGGGCCCTTTGTGCTGGGTGCGTGTCTGG	2100
Qy	2110	ACCAGACACCCAGGATGTACTTTGTTAAGCGAGATGTGACCGGGGCCCTATGATGCCATCC	2169
Db	2101	ACCAGACACCCAGGATGTACTTTGTTAAGCGAGATGTGACCGGGGCCCTATGATGCCATCC	2160
Qy	2170	CCCAGGGTAAGCTGCTGGAGGTTGTTGGCAATATGATCAGGCACATCGGAGAGCAGTACT	2229
Db	2161	CCCAGGGTAAGCTGCTGGAGGTTGTTGGCAATATGATCAGGCACATCGGAGAGCAGTACT	2220
Qy	2230	GTATCCGCCAGTATCAGTGTGTCGGAGAGATAGCCAAGGCCAAGTCCACAAGTCCCTTTA	2289
Db	2221	GTATCCGCCAGTATCAGTGTGTCGGAGAGATAGCCAAGGCCAAGTCCACAAGTCCCTTTA	2280
Qy	2290	GGAGACAGGTCAACCCCTCTCTGACCTCCAGCCATACATGGCCAGTTCCTTAAGCATC	2349
Db	2281	GGAGACAGGTCAACCCCTCTCTGACCTCCAGCCATACATGGCCAGTTCCTTAAGCATC	2340
Qy	2350	TGCAGGATTTCAGATGCCAGTGCACATGAGGAACCTCGTTGTGCATCGAGCAGCATCTCTA	2409
Db	2341	TGCAGGATTTCAGATGCCAGTGCACATGAGGAACCTCGTTGTGCATCGAGCAGCATCTCTA	2400
Qy	2410	TGAATGAGACAGCAGCAGCTGTTTGACTTCTTCTCGACTTCCTCGGTACAGGATCGG	2469
Db	2401	TGAATGAGACAGCAGCAGCTGTTTGACTTCTTCTCGACTTCCTCGGTACAGGATCGG	2460
Qy	2470	TAAAGATTGGTGACAGTGCTATACGAGTGCATGAGGAACAGCTGTTTGGTGGAGGTCC	2529
Db	2461	TAAAGATTGGTGACAGTGCTATACGAGTGCATGAGGAACAGCTGTTTGGTGGAGGTCC	2520
Qy	2530	CCACCCCTGCTCTGCAGTCTGTGTTTCGGAGACATGAGGAACAGCTGTTTGGTGGAGGTCC	2589
Db	2521	CCACCCCTGCTCTGCAGTCTGTGTTTCGGAGACATGAGGAACAGCTGTTTGGTGGAGGTCC	2580
Qy	2590	AGCGGGATGGTGGCTTTTACGTTTGTGATGACTTCTGTTGGTGGAGCCCTCACTTGG	2649
Db	2581	AGCGGGATGGTGGCTTTTACGTTTGTGATGACTTCTGTTGGTGGAGCCCTCACTTGG	2640
Qy	2650	ACCAAGCAAAAACCTTCTCAGCACCTGTGTCATGCCGTCTCTGAGTATGGGTGCATGA	2709
Db	2641	ACCAAGCAAAAACCTTCTCAGCACCTGTGTCATGCCGTCTCTGAGTATGGGTGCATGA	2700

Qy	2710	TAAACTTSCAGAAGACAGTGGTGAACCTTCCTGTGGAGCCTGGTACCCCTGGGTGGTCCAG	2769
Db	2701	TAAACTTSCAGAAGACAGTGGTGAACCTTCCTGTGGAGCCTGGTACCCCTGGGTGGTCCAG	2760
Qy	2770	CTCCATACCAGCTGCCGTGCTACATGCGCTGTTCCCTGGTGTGCTTGCCTGCTGGACACTC	2829
Db	2761	CTCCATACCAGCTGCCGTGCTACATGCGCTGTTCCCTGGTGTGCTTGCCTGCTGGACACTC	2820
Qy	2830	AGACTTTGGAGGTGTTCTGTGACTACTCAGGTTATGCCCGAGACCTCAATTAAGACGAGCC	2889
Db	2821	AGACTTTGGAGGTGTTCTGTGACTACTCAGGTTATGCCCGAGACCTCAATTAAGACGAGCC	2880
Qy	2890	TCACCTTCCAGAGTGTCTTCAAAAGCTGGGAAGACCATGGGGAACAAGCTCCCTGTCGGTCT	2949
Db	2881	TCACCTTCCAGAGTGTCTTCAAAAGCTGGGAAGACCATGGGGAACAAGCTCCCTGTCGGTCT	2940
Qy	2950	TGGGGTTGAAGTGTACGGTCTAATTTCTAGACTTCAGGTGAACAGCCCTCCAGACAGTCT	3009
Db	2941	TGGGGTTGAAGTGTACGGTCTAATTTCTAGACTTCAGGTGAACAGCCCTCCAGACAGTCT	3000
Qy	3010	GCATCAATATATACAAGATCTTCCCTGCTTCAGCCCTACAGGTTCCATGATGTGTGATTC	3069
Db	3001	GCATCAATATATACAAGATCTTCCCTGCTTCAGCCCTACAGGTTCCATGATGTGTGATTC	3060
Qy	3070	AGCTTCCCTTTGACCAGCGTGTTAGGAAGAACTCACATTCCTTCTGGGCATCATCTCCA	3129
Db	3061	AGCTTCCCTTTGACCAGCGTGTTAGGAAGAACTCACATTCCTTCTGGGCATCATCTCCA	3120
Qy	3130	GCCAAGCATCTTGCTGCTATGCTATCCTGAAGGTCAAGAATCCAGGAATGACACTAAAGG	3189
Db	3121	GCCAAGCATCTTGCTGCTATGCTATCCTGAAGGTCAAGAATCCAGGAATGACACTAAAGG	3180
Qy	3190	CCTCTGGCTCCCTTCTCCTGAAGCGGCACATTTGGCTCTGTCTACAGGCCCTTCCTGCTCA	3249
Db	3181	CCTCTGGCTCCCTTCTCCTGAAGCGGCACATTTGGCTCTGTCTACAGGCCCTTCCTGCTCA	3240
Qy	3250	AGCTGGCTGCTCATCTGCTCATCTACAATGTCTCCTGGGACCTCTGAGGACAGCCCAAA	3309
Db	3241	AGCTGGCTGCTCATCTGCTCATCTACAATGTCTCCTGGGACCTCTGAGGACAGCCCAAA	3300
Qy	3310	AACTGCTGTGGCGGAGCTCCAGAGGGCGCAATGACCATCCTTTAAAGCTGCAGCTGACC	3369
Db	3301	AACTGCTGTGGCGGAGCTCCAGAGGGCGCAATGACCATCCTTTAAAGCTGCAGCTGACC	3360
Qy	3370	CAGCCCTAAGCACAGACTTTCAGACCATTTTGGACTTAACCTGTCTCCTTCCCGCTAGATG	3429
Db	3361	CAGCCCTAAGCACAGACTTTCAGACCATTTTGGACTTAACCTGTCTCCTTCCCGCTAGATG	3420
Qy	3430	AACATG 3435	
Db	3421	AACATG 3426	

RESULT 2

AF073311	LOCUS	AF073311	3369 bp	mRNA	ROD	09-SEP-1998
	DEFINITION	Mus musculus telomerase catalytic subunit mRNA, complete cds.				
	ACCESSION	AF073311				
	NID	93551846				
	KEYWORDS	AF073311.1	GI:3551846			
	VERSION					
	SOURCE	house mouse.				
	ORGANISM	Mus musculus				
	REFERENCE	Eukaryota; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Mus.				
	AUTHORS	1 (bases 1 to 3369)				
	TITLE	Martín-Rivera-L., Herrera,E., Albar,J.P. and Blasco,M.A.				
	JOURNAL	Expression of mouse telomerase catalytic subunit in embryos and adult tissues				
	MEDLINE	Proc. Natl. Acad. Sci. U.S.A. 95 (18), 10471-10476 (1998)				
	REFERENCE	98393668				
		2 (bases 1 to 3369)				

AUTHORS Martin-Rivera L., Herrera E. and Blasco M.A.
TITLE Direct Submission
JOURNAL Submitted (19-JUN-1998) Immunology and Oncology, National Centre of Biotechnology, Cantoblanco, Madrid 28049, Spain

FEATURES

source 1. .3369
/organism="Mus musculus"
/db_xref="taxon:10090"

CDS

1. .3369
/codon_start=1
/product="telomerase catalytic subunit"
/protein_id="AAC34821.1"
/db_xref="PDB:3551847"
/translation="MTAPRCAPAVRSLLRSRYREVPWPLATFVRRLGPEGRRLVOPGDP
LNEARAGPMAFTSVRSYLPNTVETLRVSGAMWLLSRVDDLLVLLAHALYL
LVPPSAYOVCGSLYQICATTDIWPVSASACYPVPRVGEGRHROVLPFGSKSWPSP
APKALPREGKRLHSLJSTSVPSAKACYPVPRVGEGRHROVLPFGSKSWPSP
ARPEVPTAEKDLSSGKVSJSLSGSVCKKPSSTSLSPRONALORLPFTETRH
FLYSGDGGERLNPFLSLNLPNTGARRVLELIFLSRPTSGPLORHLRSLRYV
QMRPLQQLVNHAEQYVRLRSHCRFTANQVTDALNTSPHMLDLRLHSPWQ
WYGLRACILKVVSASLWGTNRHREFFKLFYSLGKYLKSLQELMMWKKVEDCH
WLRSSPGKDRVPAEHLRLERLRELSEVRRHQDTWLPICRLRFPKNGRLPIV
RKSWSKLSIGVROHLEHRLERLRELSEVRRHQDTWLPICRLRFPKNGRLPIV
NMSYSGTRALGRKQAEHTORLTLFSLNVERTKHPHLMGSSVGLMNDIYRTWRA
FVLRALDQTPRMVFKADYGCAYDAIPQGLVEVANNIRHSESTYCIROYAVVRR
DSQGVHFRVOTLSDLOPNGOIFQSGSSSLSLGSLFCFDMENKLEAEVORGLL
DFDLFLHRSYVKIGDRYTCQGIQSGSSSLSLGSLFCFDMENKLEAEVORGLL
LRFVDDPLVLPDLQAKTFLSLVHVEYGMINLQTVNVPFVPEVLTGGAAPYR
LPACHLPWGLLDLTQTEVFDYQYATSIKSLTQSFQKAGTKMRNLLSLYLR
LKCHGLFDLVQNSLQTVINIKYIFLLQAYRHACVIOIPDFDQVRYKRLTFLGLIS
SQASCCVAILKVPNGMTLAKSGFPEAAHWCYQAFLLKLAHSVIVYKLLGLPLRT
AQKLRKLEATWILKAAADPALSTDFQTLID"

BASE COUNT 737 a 959 c 891 g 782 t
ORIGIN

Query Match 96.3%; Score 3367.4; DB 12; Length 3369;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3368; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 39 ATGACCCGCGCTCTCGTGGCCGCGGCTCTCTGCTGCGCAGCCGATACCGGAG 98
Db 1 ATGACCCGCGCTCTCGTGGCCGCGGCTCTCTGCTGCGCAGCCGATACCGGAG 60
Qy 99 GTGTGCGCGCTGCAACCTTTGTGCGGCGCTTGTGCGGCGCGAGCGCGCTTGTGCA 158
Db 61 GTGTGCGCGCTGCAACCTTTGTGCGGCGCTTGTGCGGCGCGAGCGCGCTTGTGCA 120
Qy 159 CCGGGGACCCGAGATCTACCGCACTTTGTGTTGCCCAATGCTAGTGTGATGCACTGG 218
Db 121 CCGGGGACCCGAGATCTACCGCACTTTGTGTTGCCCAATGCTAGTGTGATGCACTGG 180
Qy 219 GGTCTACAGCTCCACTGCGGACCTTTCTTCCACAGGTGTCTCCCTGAAAGAGCTG 278
Db 181 GGTCTACAGCTCCACTGCGGACCTTTCTTCCACAGGTGTCTCCCTGAAAGAGCTG 240
Qy 279 GTGGCAGGCTTGTGAGAGACTCTGCGAGCGCAACGAGAGAAACGCTGTGCTTTGGC 338
Db 241 GTGGCAGGCTTGTGAGAGACTCTGCGAGCGCAACGAGAGAAACGCTGTGCTTTGGC 300
Qy 339 TTTGAGTGTCTTAACGAGGCGAGCGGCGCTCCCTTCACTAGTACGCTGGCT 398
Db 301 TTTGAGTGTCTTAACGAGGCGAGCGGCGCTCCCTTCACTAGTACGCTGGCT 360
Qy 399 AGCTACTTGCCCAACACTGTTATGAGACCTTCGCTGTGTCAGTGGTGCATGGATGCTACTG 458
Db 361 AGCTACTTGCCCAACACTGTTATGAGACCTTCGCTGTGTCAGTGGTGCATGGATGCTACTG 420
Qy 459 TTGAGCCGAGTGGCGAGGACCTGCTGCTACCTGCTGCGACACGTGCTCTTTATCTT 518
Db 421 TTGAGCCGAGTGGCGAGGACCTGCTGCTACCTGCTGCGACACGTGCTCTTTATCTT 480

Qy 519 CTGGTCCCGCCAGCTGTGCTTACCAGGTGTGGTGTCCCGCTGTACCAATTTGTGCC 578
Db 481 CTGGTCCCGCCAGCTGTGCTTACCAGGTGTGGTGTCCCGCTGTACCAATTTGTGCC 540
Qy 579 ACCACGATATCTGGCCCTCTGTGTCCCGTGTAGTACAGGCCACCGCCGCTGGGAGG 638
Db 541 ACCACGATATCTGGCCCTCTGTGTCCCGTGTAGTACAGGCCACCGCCGCTGGGAGG 600
Qy 639 AATTCTCACTAACCTTAGGTTCTTACACAGATCAAGAGCAGTAGTCGCCAGGAACACCG 698
Db 601 AATTCTCACTAACCTTAGGTTCTTACACAGATCAAGAGCAGTAGTCGCCAGGAACACCG 660
Qy 699 AAACCCCTGGCCCTTGCCATCTCGAGGTACAAAGAGCATCTGAGTCTCACCACTCAAGT 758
Db 661 AAACCCCTGGCCCTTGCCATCTCGAGGTACAAAGAGCATCTGAGTCTCACCACTCAAGT 720
Qy 759 GTGGCTTACAGTAAAGAGGCGCAGATGCTATCTCTCCGAGAGTGGAGGAGGACCCAC 818
Db 721 GTGGCTTACAGTAAAGAGGCGCAGATGCTATCTCTCTCCGAGAGTGGAGGAGGACCCAC 780
Qy 819 AGGAGGTGTGTACCAACCCCATCAGGCAAAATCATGGGTGCCAAGTCTGCTCGGTCCCC 878
Db 781 AGGAGGTGTGTACCAACCCCATCAGGCAAAATCATGGGTGCCAAGTCTGCTCGGTCCCC 840
Qy 879 GAGTGCTCTGACAGAGAAAGATTTCTCTTAAAGGAAAGTGTCTGACCTGAGTCTC 938
Db 841 GAGTGCTCTGACAGAGAAAGATTTCTCTTAAAGGAAAGTGTCTGACCTGAGTCTC 900
Qy 939 TCTGGGTGCGTGTGTGTAAACAAAGCCAGCTCCACATCTCTGCTGTACACACCCCG 998
Db 901 TCTGGGTGCGTGTGTGTAAACAAAGCCAGCTCCACATCTCTGCTGTACACACCCCG 960
Qy 999 CAAATGCTTTAGCTCAGGCCATTTATTGAGACACAGACATTTCTTTACTCCAGGGA 1058
Db 961 CAAATGCTTTAGCTCAGGCCATTTATTGAGACACAGACATTTCTTTACTCCAGGGA 1020
Qy 1059 GATGCCAAGAGCGTCTAAACCCCTCATTTCTACTCAGCAACCTCCAGCTTAAGTCTGACT 1118
Db 1021 GATGCCAAGAGCGTCTAAACCCCTCATTTCTACTCAGCAACCTCCAGCTTAAGTCTGACT 1080
Qy 1119 GGGGCGAGGAGACTGGTGGAGATCATCTTTCTGGCTCAAGGCTAGGACATCAGGACCA 1178
Db 1081 GGGGCGAGGAGACTGGTGGAGATCATCTTTCTGGCTCAAGGCTAGGACATCAGGACCA 1140
Qy 1179 CTCTGAGACACACCGCTCTATCGCTGATCTGCTGAGATCGGCGCCCTGTTCACACAG 1238
Db 1141 CTCTGAGACACACCGCTCTATCGCTGATCTGCTGAGATCGGCGCCCTGTTCACACAG 1200
Qy 1239 CTGCTGGTGAACCATGCGAGTGCCCAATATGTCAGACTCTCAGCTCAGTTCAGTGTG 1298
Db 1201 CTGCTGGTGAACCATGCGAGTGCCCAATATGTCAGACTCTCAGCTCAGTTCAGTGTG 1260
Qy 1299 CGAACAGCAACCAACAGTGACAGATGCTTGAACACAGCCACCGCCACCTCATGATGAT 1358
Db 1261 CGAACAGCAACCAACAGTGACAGATGCTTGAACACAGCCACCGCCACCTCATGATGAT 1320
Qy 1359 TTGCTCGGCTGACACAGCAGTCCCTGGCAGGTATATGTTTCTTCTGGGCTGTCTCTGC 1418
Db 1321 TTGCTCGGCTGACACAGCAGTCCCTGGCAGGTATATGTTTCTTCTGGGCTGTCTCTGC 1380
Qy 1419 AAGTGTGTCTGCTAGTCTCTGGGTACAGGCAAAATAGCGGCGCTTCTTTAAGAAC 1478
Db 1381 AAGTGTGTCTGCTAGTCTCTGGGTACAGGCAAAATAGCGGCGCTTCTTTAAGAAC 1440
Qy 1479 TTAAGAGTTCATCTCGTTGGGAAATAGCGCAAGCTATFACCTGAGGAAGTGTGG 1538
Db 1441 TTAAGAGTTCATCTCGTTGGGAAATAGCGCAAGCTATFACCTGAGGAAGTGTGG 1500
Qy 1539 AAGTGAAGTAGAGGATTCGCCACTGCTCCGAGGAGCCCGGGAGGACCGGTGTCCCC 1598
Db 1501 AAGTGAAGTAGAGGATTCGCCACTGCTCCGAGGAGCCCGGGAGGACCGGTGTCCCC 1560

AUTHORS Counter,C.M., Meyerson,M., Eaton,E.N., Ellisen,L.W., Caddle,S.D., Haber,D. and Weinberg,R.A.
TITLE Direct Submission
JOURNAL Submitted (20-JAN-1998) Whitehead Institute, 9 Cambridge Center, Cambridge, MA 02142, USA
FEATURES Location/Qualifiers
1..8960
/organism="synthetic construct"
/db_xref="taxon:32630"
order(1..1101,4606..8960)
/note="mammalian expression vector; see GenBank Accession Number U47120"
/organism="Cloning vector pCI-neo"
/db_xref="taxon:45772"
102..4542
/note="telomerase catalytic subunit hTERT/hEST2 mRNA"
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="5"
/map="5p15.33"
1102..4542
/gene="hTERT/hEST2"
1147..4605
/note="telomerase/reverse transcriptase"
/codon_start=1
/transl_table=11
/product="telomerase catalytic subunit"
/protein_id="AAC40212.1"
/db_xref="PID:g3023055"
/db_xref="GI:3023055"
/translation="MFRAPRCRAVRSLLSHRYEVLPLATFVRRLPGQWRLVGRGP
AAFFALVAQCLVCPMDRAPPPAASFQVSCLEKELVARLQRLCERGAKNVLAFFA
LLDARGGPEAFETTSVRSYLPNTVDALRGSGALLLRVGDVLLHLLARCALFV
LVAPSCAYQVGPPLYLQGAATQARPPPHASGPRRLGCRERAWNSVREAGVPLGLPA
PGARRGSGASRLPLPKPRREGAPEPEPTVGSGWAHPGRTGSPDRGFCVVPSP
PAAEATSLGALSGLRHSFVSQROHAGPSTSRPPMDPTCPPIVAETKHFYLS
SGDEQLRPSGLLSRLPSLTGARLAVETIFLGRPMPTPRRLPRLPQRNRL
FLELGHNAQCPVGLLTHCPRLAAVTTPAAGVCAAREKPGQSVAAPEEDDPRKLQ
LLRQHSFGCAWLRSPGCVPAAREHRLREILAKFLHLMSSVYVYVLLRSFYVET
TKMSVSGCAWLRSPGCVPAAREHRLREILAKFLHLMSSVYVYVLLRSFYVET
TFQNRLLFFYKRSWSKLSIGIRHKLKRVQLRSEAEVROHREARPLLSRLFI
PKPGLRPIVNDYVVGARTFRREKRAERTSVKVDYDTIPDRLTEVIAIIPKNTYC
LDDTHRAVTLVRAQDPPPELYFVKVDYDTIPDRLTEVIAIIPKNTYC
VRRVAVQKAHGHVRAKFSVSTLDLQYMRQFVAHQETSPURDVAVIEQSSLL
NEASGGLDFLRMCHAVRIKGSVYQCGQPGQISLTLCSLCYGDMEKFLFAG
IRDGLLRVDDLLPHLTHAKTFLRVLVRSYVYVLLRSFYVET
GGAFTVQMPAGLFPWGLLDDTLFVSDYSYARTSIRASITFRNFGKAGRNMR
KLFGVLRKLSLILKAKNAGSLGKAGNPLPSEAVQWLCHQAFLLKLRHRT
VYPLGLSLRTAQTLRSKLPTTLTALAANPALPSDFKTLIDYVPDYAGPYD
VPDYAG"
4543..4605
/note="engineered based on the epitope for the
anti-influenza virus hemagglutinin monoclonal antibody
12CA5; similar to part of the Influenza A hemagglutinin
gene sequence in GenBank Accession Number M29257"
/organism="synthetic construct"
/db_xref="taxon:32630"
1886 a 2551 c 2484 g 2038 t 1 others
BASE COUNT 1886 a 2551 c 2484 g 2038 t 1 others
ORIGIN
Query Match 45.98; Score 1603.6; DB 15; Length 8960;
Best Local Similarity 68.8; Pred. No. 0;
Matches 2386; Conservative 0; Mismatches 974; Indels 108; Gaps 9;
QY 23 CCCGCCCTTGAGCAATGACCGCGCTCTCTGTTGCCCGCGGTGCGCTCTCTGCG 82
hh 1131 CCGGCCACCCCGCGATGCGCGGCTCCCGCTGCGAGCGCGTGGCTCTCTGCG 1190
QY 83 CAGCCGATACCGGAGGTGTGGCGCTGGCAACCTTTGTGGCGCGCTGGCGCCCGAGGG 142
Db 1191 CAGCCACTACCGGAGGTGCTGGCGCTGGCCACCGTTCGTGGCGCGCTGGCGCCCGAGGG 1250

QY 143 CAGCGCCTTGTGCAACCCGGGACCCGAGATCTACCGCACTTTGGTTGCCAATGCCT 202
Db 1251 CTGCGGCTGTGTGAGCGCGGACCCGCGGCTTTCCGCGGCTGGTGGCCAGTGCT 1310
QY 203 AGTGTGATGCACTGGGGCTCAGCGCTCCACTGCGGACCTTCTCTCCACGAGTGTC 262
Db 1311 GGTGTGCTGCGCTGGGACGACGGCGCCCGCGGCGCTCTCTCCGCGAGTGTC 1370
QY 263 ATCCCTGAAGAGCTGTGGCGGAGGTGTGAGAGACTCTGCGAGCGCAAGAGAA 322
Db 1371 CTGCTGAAGAGCTGTGGCGGAGGTGTGAGAGCTGTGCGAGCGCGGCGCAAGAA 1430
QY 323 CGTGTGCTGCTTTGGCTTTGAGCTGTGCTTAAGAGGCGAGGCGGCTCTCCATGCGCT 382
Db 1431 CGTGTGCTGCTTTGGCTTTGAGCTGTGCTTAAGAGGCGAGGCGGCTCTCCATGCGCT 1490
QY 383 CACTGTAGCTGTGGCTAGCTACTTGTCCCAACACTGTATTGAGACCTGTGTGAGTGG 442
Db 1491 CACCACCGAGCTGTGGCGAGCTACTTGTCCCAACACTGTATTGAGACCTGTGTGAGTGG 1550
QY 443 TGCATGATGCTACTTGTGAGCGGAGTGGGCGAGCACTGTGCTGTCTACCTGTGGCACA 502
Db 1551 GCGGTGGGCTGCTGTGCGCGCGTGGGCGAGCACTGTGCTGTCTACCTGTGGCACA 1610
QY 503 CTGTGTCTTTATCTGTGCTGCCCGCGAGCTGTGCTTACCAGGTGTGTGGGTCTCCCT 562
Db 1611 CTGCGGCTCTTTGTGTGCTGTGCTGCCAGCTGTGCTTACCAGGTGTGTGGGTCTCCCT 1670
QY 563 GTACCAAAATTTGCGCACCAACGAGTATCTGGCCCTCTGTGCTGCTGTCTACAGGCCAC 622
Db 1671 GTACCAAGTGGCGCTGCCACTAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1730
QY 623 CCGACCGTGGCGAGGAAATTTCACTAACCTTAGTGTCTTACCAAGATCAAGAGCAGTAG 682
Db 1731 CGCTCTGGGATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1763
QY 683 TCGCGAGGAGGACCGCAACCCCTGCGCTTGCATCTCAGGTACAAAGAGGCACTGTAG 742
Db 1764 CGTCAGGAGGCGCGGCTGCGCTTGCATCTCAGGTACAAAGAGGCACTGTAG 802
QY 743 TCTCAGCAGTACAGTGTGCTTGCATCTCAGGTACAAAGAGGCACTGTAG 883
Db 1824 CAGTGTCCAGCGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1883
QY 803 GGAGGAGGAGGACCGCAACCCCTGCGCTTGCATCTCAGGTACAAAGAGGCACTGTAG 862
Db 1884 GGAGGAGGAGGACCGCAACCCCTGCGCTTGCATCTCAGGTACAAAGAGGCACTGTAG 1943
QY 863 TCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 913
Db 1944 TGACCGTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2003
QY 914 AGGAAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 970
Db 2004 GGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2063
QY 971 CTCACAT---CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1027
Db 2064 CCCCCCATCCATCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2123
QY 1028 TGAGACGACAGCTTCTTCTTACTTCCAGGAGATGCGCAAGAGGCTCTAAACCCCTCAT 1087
Db 2124 CGAGACCAAGCACTTCTCTTACTTCTCTCAGGCGA---CAAGGAGGAGCTGCGGCGCTCT 2180
QY 1088 CTTACTCAGCAACTCCAGCCTTAACCTTGTGCTGGGCGAGGAGCTGGTGGAGATCATCTT 1147
Db 2181 CTTACTCAGCTCTCTGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2240
QY 1148 TCTGGGCTCAAGGCTTAGGACATCAGGACCACTCTGCGAGGACACCGCTCTATCGCGTCG 1207
Db 2241 TCTGGTTCAGGCGCTGGATGCGAGGAGCTCCCGCGAGGCTGCGCGCGCTGCGCGCGCGCG 2300

Db 1033 CGAGACCAAGCACTTCTCTACTCTCAGGCGA---CAAGGAGCAGCTGCGGCCCTCCTT 1089
Qy 1088 CCTACTCAGCAACCTCCAGCCTAACTTGAATGGGCGCAGGAGCTGTGGAGATCATCTT 1147
Db 1090 CCTACTCAGCTCTCTGAGGCCAGCCTGACTGCGCTCGAGGCTGTGGAGACCATCTT 1149
Qy 1148 TCTGGGCTCAAGGCTTAGACATCAGGACCACTCTGAGGACACACCGCTATCGGGTGG 1207
Db 1150 TCTGGGCTTCCAGGCTTGATGCCAGGACTTCCCGGAGGTTGCGCGCTGCCCCAGCG 1209
Qy 1208 ATACTGGCAGATGCGGCCCTTGTCCAAACAGCTCTGCTGAACATGACAGATGCCAATA 1267
Db 1210 CTACTGGCAATGCGGCCCTTGTCTGGAGCTGCTTGGAAACAGCGGAGTGGCCCTA 1269
Qy 1268 TGTAGACTCTCAGGTCACATTCAGGTTTCCAAACAGCAACCAACAGGTGACAGATGC 1327
Db 1270 CGGGTGTCTCAGAGCGCACTGCCGCTGCGAGTGTGGGTGACCCAGCAGCGGGTGT 1329
Qy 1328 CT-----TGACACACAGCCC 1342
Db 1330 CTGTGCGCGGAGAACCCAGGGCTCTGTGCGGCCCGCCAGGAGGAGACAGACCC 1389
Qy 1343 ACCGCACTCATGATTTGCTCGCCTGCACAGCAGTCCCTGGCAGGTATATGTTTTCT 1402
Db 1390 CGTGGCTGTGACGTGCTCGCAGCAGCAGAGCCCTGGCAGGTGTACGGCTTCT 1449
Qy 1403 TCGGGCTGTCTCTGAAGGTGGTGTGTGCTAGTCTCTGGGTACACAGGACATGAGCG 1462
Db 1450 GCGGGCTGCTCGCGGCTGGTCCCGCCAGGCTCTGGGGCTCCAGGACCAACGAGC 1509
Qy 1463 CGCTCTTTAAGACTTAAAGAGTTCATCTGCTGGGGAATACGGCAAGCTATCACT 1522
Db 1510 CGCTTCTCAGGAACACAAAGAGTTCATCTCCTGGGGAAGCATGCCAAGCTCTCGCT 1569
Qy 1523 GCAGAACTGATGTGGAAGATGAAGTAGAGGATGCACTGGCTCCGAGGAGAGTCCCTGGCTACGTCT 1582
Db 1570 GCAGAGCTGACGTGGAGATGACGTGCGGAGTCTGCTGCTGCGCAGGAGCCAGG 1629
Qy 1583 GAAGACCGTGTCCCGCTCGAGGACACCTCTGAGGAGAGAGTCCCTGGCTACGTCTCT 1642
Db 1630 GTTGGCTGTGTTCCGCGCCAGAGACCGTCTGCTGAGGAGATCTCTGGCAAGTCTCT 1689
Qy 1643 GTTGGCTGATGGACACATACCTGTACAGTCTGTAGGTCATTTTACATCAGCA 1702
Db 1690 GCACTGCTGATGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1749
Qy 1703 GAGCACAATTCAGAAACAGGCTCTTCTACCTGAGGAGTGTAGGTCATTTTACATCAGCA 1762
Db 1750 GACCAGTTTCAAAGAACAGGCTCTTCTACCGGAGAGTGTCTGAGGAGCAAGTTGCA 1809
Qy 1763 GACATTGGAGTCAGGCAACACTTGACAGAGTGGGCTACGGAGGCTGTCAACAGGAG 1822
Db 1810 AAGCATTGAATCAGACAGCACTTGAAGAGGTCAGCTGCGGAGTGTCTGGAAGCAGA 1869
Qy 1823 GGTAGGATACCAAGGACCTGGTAGCCATGCCATCTGAGACTGCGGCTTATAGATGGTACCAAGCTTT 1942
Db 1870 GGTAGGACGATCGGGAAGCAGGCGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1989
Qy 1883 CAAGCCCAACGCTGCGGCCCTTGTGAACATGATGATGATGATGATGATGATGATGATGAT 2002
Db 1930 CAAGCCTGACGGCTGCGCGGATGTGAACATGATGATGATGATGATGATGATGATGAT 2049
Qy 1943 GGGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2062
Db 1990 CCGCAGAGAAAGAGGCGCGGAGGCTACCTCGAGGGTGAAGGACATGTTTACGGTGT 2109
Qy 2003 CAATATGACCGGACAAAATCCTCACCTTATGGGTCTTCTGTACTGGGTATGAATGA 2122
Db 2050 CAATAGAGGCGCGCGCGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2169
Qy 2063 CATCTACAGGACTGCGGGCTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2182
Db 2110 TATCCAGGCGCTGGCGACCTTCGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2199

Qy 2123 GATGTACTTTTAAAGCAGATGTGACCGGGCCCTATGATGCCATCCCCAGGTAAGCT 2182
Db 2170 GCTGTACTTTGTCAAGTGTGACGGGCGGTACGACACCATCCCCAGGACAGGCT 2229
Qy 2183 GGTGGAGGTTTGGCAATATATGATCAGGCACTCGGAGAGCAGTACTGTATCCGCCAGTA 2242
Db 2230 CAGGAGGTCATCGGCAGCATCATCA---AACCCAGAACACAGTACTGCTGGTGGGTA 2286
Qy 2243 TGCAGTGGTCCGAGAGATAGCAGGCCCAAGTCCCAAGTCTCTTTAGGAGACAGTGTAC 2302
Db 2287 TGCCTGTGTCAGAGAGCGGCCCATGGGACAGTCCGCAAGGCTTCAAGAGCCAGCTCTC 2346
Qy 2303 CACCCTCTCTGACCTCCAGCCATACATGGCCAGTTCCTTTAAGCATCTCGAGATTCAGA 2362
Db 2347 TACCTTGACAGACCTCCAGCGTACATGCGACAGTTCTGGTCTCAGCTGCAG-----GA 2400
Qy 2363 TGCCAGTGCATGAGGAATCCGTTGTCATCGAGAGCAGTCTCTATGAATGAGAGCAG 2422
Db 2401 GACCAGCCGCTGAGGATGCCGTCGTCATCGAGCAGAGCTCCTCCTGAATGAGGCGAG 2460
Qy 2423 CAGCAGCTCTGTTGACTTCTCTGCACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2482
Db 2461 CAGTGGCTCTTCCGAGCTTCTCTAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2520
Qy 2483 CAGTGTCTATACGAGTGCAGGATCCCGCAGGCTTCCAGGCTTCCAGCTTCCAGCTGCTG 2542
Db 2521 CAAAGTCTTACGTCAGTGCAGGGATCCCGAGGGTCCATCTCTCCAGCTGCTGCTG 2580
Qy 2543 CAGTCTGTGTTTCGAGAGATGAGAAACAGCTGTTTGTGCTGAGTGCAGGCGGATGGGTT 2602
Db 2581 CAGCTGTGCTACGCGGACATGGAGAAAGCTGTTTGGCGGGATTCGGCGGAGCGGCT 2640
Qy 2603 GCTTTTACGTTTGTGATGACTTCTTGTGTGACGCTTCTTGGACACCTTCCAGCAAGCAAAAC 2662
Db 2641 GCTCTGCTGTTGTTGGATGATTTCTTGTGTGACACCTTCCAGCTTCCAGCGGAGGCT 2700
Qy 2663 CTCTCTACGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2722
Db 2701 CTCTCTACGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2760
Qy 2723 GACAGTGGTGAACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2782
Db 2761 GACAGTGGTGAACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2820
Qy 2783 GCTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2842
Db 2821 GCGGGCCAGCGGCTATCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2880
Qy 2843 GTTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2902
Db 2881 GCAGAGGCTACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2940
Qy 2903 TGTCTTAAAGCTGGGAAGACCATGCGGACACAGCTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2962
Db 2941 CGGCTTCAAGCTGGGAGAACATGCGTTCGCAAACTTTTGGGGTCTTGGGCTGCTGCTGCTGCTGCT 3000
Qy 2963 TCAGGCTCTATTTCTAGACTTGCAGTGAACAGCTTCCAGAGCTTCCAGAGGCTGCTGCTGCTGCTGCT 3022
Db 3001 TCAGAGCTGTTTCTGATTTGAGGTTGAACAGCTTCCAGAGGCTTCCAGAGGCTGCTGCTGCTGCTGCT 3060
Qy 3023 CAAGATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3082
Db 3061 CAAGATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3120
Qy 3083 CCAGCGTGTAGGAAGACCTCAGATCTTCTGGGCTATCATCTCCAGGCTGCTGCTGCTGCTGCTGCTGCT 3142
Db 3121 TCAGCAAGTTTGAAGAACCCACATTTTCTGCGCTCATCTGACAGGCTGCTGCTGCTGCTGCTGCTGCT 3180
Qy 3143 CTGCTATGCTATCTGAGGCTCAAGATCCAGGATGACACTAAG-----GCCTC 3193
Db 3181 CTGCTACTCCATCTTGAAGCAAGACGAGGATGCTGCTGGGGCCCAAGGGCGCCG 3240

us-09-042-460-1.1.rge

Mon Aug 16 10:21:24 1999

```
/db_xref="PID:q4580665"
/db_xref="GI:4580665"
/translation="MTRAPCPAVRSLLRSYREVLPATVRRRLPGEGRRLVQPGDP
KIYRTIQAOLVCMWGSQPPADLSHOVSLLKELVAVRQLCERNERNVLAFGE
LLNEARGPPMAFTSVRSIYLPNVIETLVSGAMLLSRVGDVLLYLALHACALY
LVPPCAYQVCGSPYQICATDIPWSVASVYRTPVPGPHROVLFTPGKSWPSP
APKPLAPSRGTRKRLSLSTSVPSAKKARCVPVVEGPHROVLFTPGKSWPSP
ARSPVPTAKDLSSKGVSDLSLGSVCCSKHPSSTLLSPRONAFQLRPFIERH
FLYSGDGOERLNPFLSNLQNLGARRLVEIIFGSRPTSGPLCRTRHLSRPWQ
OMRPLFOQLLVNHAECQYVRLRSHCRFTANQOVTDALNTSPHMLDLRLHSSPWQ
VYGFRLACLKCVYSASLWGRNRRFRFNKLFISLKYKLSLQELMWMKVEDCH
WLRSSP"
BASE COUNT      1944 a 1857 c 1815 g 1882 t
ORIGIN

Query Match      41.8%; Score 1462.6; DB 12; Length 7498;
Best Local Similarity 93.2%; Pred. No. 0;
Matches 1583; Conservative 0; Mismatches 14; Indels 102; Gaps 1;

QY 10 GTGGAGGCCCATCCGGCCTTGAGCACAATGACCCGCTCTCTCTGTTGGCCCGCGTGC 69
DB 4458 GTGGAGGCCCATCCGGCCTTGAGCACAATGACCCGCTCTCTCTGTTGGCCCGCGTGC 4517
QY 70 GCTCTCTGCTGCGCAGCGGATACCGGGAGGTGTGGCGCTGCGAACCTTTGTGGGGGCC 129
DB 4518 GCTCTCTGCTGCGCAGCGGATACCGGGAGGTGTGGCGCTGCGAACCTTTGTGGGGGCC 4577
QY 130 TGGGGCCCGAGGAGCGGCTTGTGCAACCGGGAGCCCGAAGATCTACCGCACTTTGG 189
DB 4578 TGGGGCCCGAGGAGCGGCTTGTGCAACCGGGAGCCCGAAGATCTACCGCACTTTGG 4637
QY 190 TTGCCCAATGCTAGTGTGCATGCACTGGGCTCAGAGCTCCACCTCGCGACCTTTCT 249
DB 4638 TTGCCCAATGCTAGTGTGCATGCACTGGGCTCAGAGCTCCACCTCGCGACCTTTCT 4697
QY 250 TCC----- 253
DB 4698 TCCACCAAGTGGGCTTCCAGCGGGATCCCATGGTGCAGGGCGGAAACCGGGAGGAC 4757
QY 253 -----ACCAAGTGTATCTCC 267
DB 4758 GTGGATAGTGGCTAGCTCATGTCTAAGACCTCTTCTCTTACCAAGTGTATCTCC 4817
QY 268 TGAAGAGCTGTGGCCAGGTTGTGCAGAGACTGTGGAGCGGCAACGAGAGAACGTGC 327
DB 4818 TGAAGAGCTGTGGCCAGGTTGTGCAGAGACTGTGGAGCGGCAACGAGAGAACGTGC 4877
QY 328 TGGCTTTTGGCTTTGAGCTGCTTAAGAGCCAGAGCGGGCTCCCATGGCTTCACTA 387
DB 4878 TGGCTTTTGGCTTTGAGCTGCTTAAGAGCCAGAGCGGGCTCCCATGGCTTCACTA 4937
QY 388 GTAGCGTGTAGCTACTTGTGCCAACACTGTATTGAGACCTTGTGGTGTGAGTGTGCAT 447
DB 4938 GTAGCGTGTAGCTACTTGTGCCAACACTGTATTGAGACCTTGTGGTGTGAGTGTGCAT 4997
QY 448 GGATGCTACTTGTAGCGAGTGGCGACGACCTGTGGTGTGAGTGTGAGTGTGAGTGTG 507
DB 4998 GGATGCTACTTGTAGCGAGTGGCGACGACCTGTGGTGTGAGTGTGAGTGTGAGTGTG 5057
QY 508 CTCCTTTATCTTCTGGTGGCCCGGAGTGTGCTACCAAGTGTGTGGTGTGCTCCCTGTACC 567
DB 5058 CTCCTTTATCTTCTGGTGGCCCGGAGTGTGCTACCAAGTGTGTGGTGTGCTCCCTGTACC 5117
QY 568 AAATTGTGCCACACGAGATATCTGGCCCTCTGTGTGGTGTGAGTGTGAGTGTGAGTGTG 627
DB 5118 AAATTGTGCCACACGAGATATCTGGCCCTCTGTGTGGTGTGAGTGTGAGTGTGAGTGTG 5177
QY 628 CCGTGGGAGGAATTCTACCTTAGTGTCTTACACAGATCAAGAGCAGTAGTGTGCC 687
DB 5178 CCGTGGGAGGAATTCTACCTTAGTGTCTTACACAGATCAAGAGCAGTAGTGTGCC 5237
QY 688 AGGAGCACCGAAACCCCTGGCCTTGGCCTTGGCCTTGGCCTTGGCCTTGGCCTTGGCCTTCA 747
```

```
Db 5238 AGGAAGCACCAGAACCCCTGGCCCTTCGCACTCGAGGTACAAAGAGGCATCTGAGTCTCA 5297
Qy 748 CCAGTACAAGTGTGCTTCAGTAAGAAGGCCAGATGCTATCCTGTCCGAGAGTGAGG 807
Db 5298 CCAGTACAAGTGTGCTTCAGTAAGAAGGCCAGATGCTATCCTGTCCGAGAGTGAGG 807
Qy 808 AGGGACCCACAGGAGGTGTCTACCAACCCCATCAGCAATCATGCTGCTCCGAGAGTGAGG 5357
Db 5358 AGGGACCCACAGGAGGTGTCTACCAACCCCATCAGCAATCATGCTGCTCCGAGAGTGAGG 867
Qy 868 CTCGCTCCCGAGGTGCTCTACTCAGAGAAAGATTGTCTCTTAAGGAAAGGTGCTG 5417
Db 5418 CTCGCTCCCGAGGTGCTCTACTCAGAGAAAGATTGTCTCTTAAGGAAAGGTGCTG 927
Qy 928 ACCTGAGTCTCTCTGGGTGGGTGCTGTAAACACAGCCAGCCAGCTCCACATCTCTGCTGT 5477
Db 5478 ACCTGAGTCTCTCTGGGTGGGTGCTGTAAACACAGCCAGCCAGCTCCACATCTCTGCTGT 987
Qy 988 CACCACCCGCCAAATGCTTTTCACTCAGGCGCATTTTATGAGACCAACATCTCTGCTGT 5537
Db 5538 CACCACCCGCCAAATGCTTTTCACTCAGGCGCATTTTATGAGACCAACATCTCTGCTGT 1047
Qy 1048 ACTCCAGGGAGATGGCCAGAGCGCTTAACCCCTCATCTCTACTCAGCAACCTCCAGC 5597
Db 5598 ACTCCAGGGAGATGGCCAGAGCGCTTAACCCCTCATCTCTACTCAGCAACCTCCAGC 1107
Qy 1108 CTAACCTGACTGGGCGCAGGAGTGGTGGAGATCATCTTCTGGGCTCAAGGCGCTAGGA 5657
Db 5658 CTAACCTGACTGGGCGCAGGAGTGGTGGAGATCATCTTCTGGGCTCAAGGCGCTAGGA 1167
Qy 1168 CATCAGGACCACTCTGCAGGACACACCGCTCTATCGCTGATATGCGAGATGCGGCGCC 5717
Db 5718 CATCAGGACCACTCTGCAGGACACACCGCTCTATCGCTGATATGCGAGATGCGGCGCC 1227
Qy 1228 TGTTCACACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5777
Db 5778 TGTTCACACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1287
Qy 1288 ATTGAGGTTTTCGACAGCAACCAACAGGTGACAGATGCTGCTGCTGCTGCTGCTGCTGCTG 5837
Db 5838 ATTGAGGTTTTCGACAGCAACCAACAGGTGACAGATGCTGCTGCTGCTGCTGCTGCTGCTG 1347
Qy 1348 ACCTCATGATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5897
Db 5898 ACCTCATGATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1407
Qy 1408 CCTGCTCTGCAAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5957
Db 5958 CCTGCTCTGCAAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1467
Qy 1468 TCTTTAAGAACTTAAAGAAAGTTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6017
Db 6018 TCTTTAAGAACTTAAAGAAAGTTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1527
Qy 1528 AACTGATGTGAAGATGAAGTAGAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6077
Db 6078 AACTGATGTGAAGATGAAGTAGAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1587
Qy 1588 ACGGTGCTCCCGCTGCAGA 1606
Db 6138 ATGGCTGGTCTCCAGTCA 6156
```

RESULT 7

```
AF114847
LOCUS AF114847 12213 bp DNA PRI 28-FEB-1999
DEFINITION Homo sapiens telomerase reverse transcriptase (TEXT) gene, partial
ACCESSION AF114847
NID 94324972
VERSION AF114847.1 GI:4324972
KEYWORDS
```

```
SOURCE human.
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 12213)
AUTHORS Wu,K.J., Grandori,C., Anacker,M., Simon-Vermot,N., Polack,A.,
Lingner,J. and Dalla-Favera,R.
TITLE Direct activation of TERT transcription by c-MYC
MEDLINE Nat. Genet. 21 (2), 220-224 (1999)
REFERENCE 99140777
AUTHORS 2 (bases 1 to 12213)
Ducrest,A.-L., Anacker,M., Reichenbach,P., Nabholz,M. and
Lingner,J.
TITLE Identification of the human telomerase reverse transcriptase
(htERT) promoter
JOURNAL Unpublished
AUTHORS 3 (bases 1 to 12213)
Anacker,M., Simon-Vermot,N. and Lingner,J.
TITLE Direct Submission
JOURNAL Submitted (17-DEC-1998) Telomerase Research Unit, ISREC, Chemin de
Boveresses 155, Eplinges 1066, Switzerland
FEATURES
source
1. 12213
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="5"
2. 158
/rpt_family="L1"
/rpt_type="dispersed"
167..376
/rpt_family="L1"
/rpt_type="dispersed"
384..685
/rpt_family="Alu"
/rpt_type="dispersed"
764..986
/rpt_family="L1"
/rpt_type="dispersed"
3091..3345
/rpt_family="L1"
/rpt_type="dispersed"
3352..3696
/rpt_family="L1"
/rpt_type="dispersed"
3911..4293
/rpt_family="L1"
/rpt_type="dispersed"
4487..4787
/rpt_family="L1"
/rpt_type="dispersed"
5270..5582
/rpt_family="Alu"
/rpt_type="dispersed"
7453..79129
/gene="TERT"
join(7453..7671,7776..79129)
/gene="TERT"
/codon_start=1
/product="telomerase reverse transcriptase"
/protein_id="AADI7210.1"
/db_xref="PID:94324973"
/db_xref="GI:4324973"
/translation="MPAPRCRAVRSLLRSHRYREVLPATFVRRLLPGQWRLVORDDP
AAFRALVAQCLVCFWMDARPPAAPSPFQVSCLELVARVLRORLCERGAKNVLAGFA
LLDGAGGPFPAFTTSVRSYLPNTVTDALRSGAWGLLRVCGDDVLYHLARCALFY
LVAPSCATVCGPPLYQLGATQAPPPHAGSPRRRLGGERANNNHVSREAGVPLGLPA
PGARRGGSASRLPLKPRRGAAPPEPTPVQGSWAHPGTRGSDRGFCVVSFA
RPAEATSLGASLPLKPRRGAAPPEPTPVQGSWAHPGTRGSDRGFCVVSFA
SGDKEQLRPSFLLSLRPLTGARRLVETIFLGSRRMPEPTPRRLPQRYWOMRPL
FLELLGHAOCQGVLLKTHCPLRAAVTPAAGVCAAREKPGQSVAAPEEDTPRLVQ
TLROHSSPMQVYGFVRACLRRLVPPGLWGRSHNRERFLNKKKISLGHKAKLSLQEL
TWKMSVRDCAWLRSP"
7672..7775
intron
```

Mon Aug 16 10:21:24 1999

```

/gene="TERT"
/number=1
repeat_region 9629..9923
/rpt_family="Alu"
/rpt_type=dispersed
repeat_region 10314..10475
/rpt_family="L1"
/rpt_type=dispersed
repeat_region 11835..11944
/rpt_family="L1"
/rpt_type=dispersed
repeat_region 11987..12213
/rpt_family="L1"
/rpt_type=dispersed
BASE COUNT 2447 a 3436 c 3481 g 2849 t
ORIGIN
Query Match 13.48; Score 468.8; DB 42; Length 12213;
Best Local Similarity 59.48; Pred. No. 7.4e-115;
Matches 1046; Conservative 0; Mismatches 522; Indels 194; Gaps 7;
QY 23 CCCGGCTTGTGACACAAATGACCGGGCTCTCTGTCGCCCGGTGCGCTCTCTGCTGGG 82
D 7437 CCCGGCACCCCGGATGCGCGGCTCCCGCTGCGGAGCGCTCCCTGCTGGG 7496
QY 83 CAGCGATACCGGAGGTGTGCGCGCTGGCAACCTTTGTGCGGCGCTGGGCGCCGAGGG 142
D 7497 CAGCCACTACCGGAGGTGCTCCCGCTGCCACGTTCTGTCGGCGCTGGGCGCCGAGGG 7556
QY 143 CAGCGCGCTTGTGCAACCGGAGCCCGGAGATCTACCGCACTTTGTTGGTGGCCAAATGCT 202
D 7557 CTGGCGCTGTGTGACGGGGGACCCCGCGCTTTCGCGCGCTGTGGTGGCCAGTGCCT 7616
QY 203 AGTGTGATCACTGGGCTACAGCTTCCACCTGCGGACCTTTCTCCAC----- 255
D 7617 GGTGTGGCTGCCGTGGACGACGCGCGCCCGCGCGCCCTCTCTCCGCGAGGTGGG 7676
QY 255 ----- 255
D 7677 CTCCCGGGTCCGGCTCCGGTGTAGGGGCGCGGGGGAACACGACATGC 7736
QY 255 -----CAGGTGTCATCCCTGAACAGCTG 278
D 7737 GGAGAGCAGCAGCGGACTCAGGGCGCTTCCCGCAGGTGCTCCGCTGAAGAGCTG 7796
QY 279 GTGGCCAGGGTTGTCAGAGACTCTGGAGCGCAACAGAGAAACGTGCTGTTTGGC 338
D 7797 GTGGCCAGGTGCTCAGAGGCTGTGCGAGCGCGCGCGAGAGAGTGTGCTGGCTTCGGC 7856
QY 339 TTTGAGCTGTTAACAGGCGCAGAGCGGGGCTCCCATGGCTTCATAGTACGTGCGT 398
D 7857 TTGCGCTGCTGGACGGGCGCGGGGGCGCCCGAGGCTTTCACCAACAGCGTGGCG 7916
QY 399 AGCTACTTGGCCAACTGTTATTGAGACCTGCTGTCAGTGTGTCATGTCATGCTACTG 458
D 7917 AGCTACTGCTGCAACAGTGTACCGACGACGACGCTGCGGGGAGCGGGCGTGGGGCTG 7976
QY 459 TTGAGCGGAGTGGGCGACGACCTGCTGTGTCATCTGTGGGCGACAGTGTCTTTATCTT 518
D 7977 CTGGCGCGCTGGGCGACGACGCTGCTGTTTACCTGCTGGGCGCTGCGCGCTTTTGTG 8036
QY 519 CTGTGCCCCCGAGCTGCTACAGGTGTGTGGGTCTCCCTCTGCTACCAATTTGTGCC 578
D 8037 CTGTGGCTCCAGCTGGCTTACCAGGTGTGGGGCGCGCTGTACCACTGCGGCGCT 8096
QY 579 ACCACGATATCTGGCCCTGTGTGCTGTGTGTCAGTGTAGTACAGCCCGCGCTGGCGAG 638
D 8097 CCCTCAGGCGCGCGCGCGCGCTAGTGTGAGCCCGCGAGGCTGTGGATGCT--- 8154
QY 639 AATTCTACTACCTTAGTCTTCAACAGATCAAGACAGTGTGTCGCGAGGAGGACCGG 698
D 8154 -----GACGGGCTGGAACCATAGCTCAGGGAGGCGGG 8189
```

```

QY 699 AAACCCCTGGCCTTGCATCTCGAGGTACAAAGAGGCATCTGAGTCTCACAGTACAAAGT 758
D 8190 GTCCCTCTGGGCTTGCAGCCCGGCTGCAGGAGCGCGGGGAGTGCACAGCGAAAGT 8249
QY 759 GTGCTTTCAGCTAAGAAGGCGCATGCTATCTCTGCCAGAGTGGAGGAGGAGGAGCCAC 818
D 8250 CTGCGTTTGCCTAAGAGGCGCGCTGCGCTGCGCCCTGAGCCGAGGAGGAGGAGCGGTT 8309
QY 819 AGGAGGTGCTACCAACCCCATCAGCAAAATCATGGGTGCCAAGTCTCTGCTGGTCCCC 878
D 8310 GGGGAGGGTCTCGGGCCCGCCCGGAGGAGCGGTGGACCGAGTACCGTGTCTGT 8369
QY 879 GAGGTGCTCTACT-----GCAGAGAAAGATTGTTCTTAAAGAAAGAGTGTCTGAC 929
D 8370 GTGGTGTACCTGCCAGACCGCGCGGAAAGCCACCTCTTTGGAGGTGCGCTCTCTGGC 8429
QY 930 CTGAGTCTCTC---TGGGTGCGGTGCTGTAAACACAGCCAGCTCCACAT---CTCTG 983
D 8430 ACGGCGCACTCCCAACCCATCCGTGGCGCGCAGCACACCGAGCGCCCGCCCATCACATCG 8489
QY 984 CTGTCTACACCGCGCGCAAAATGCTTTCAGCTCAGGCCATTTATTGAGACGAGATTTTC 1043
D 8490 CGGCGCACCGCTCCCTGGGACACGCGCTTCCCGCGGTGTACGCGGAGACCAAGCACTTC 8549
QY 1044 CTTTACTTCAGGGAGATGGCCAAAGAGCGTCTAAACCCCTATCTCTACTCAGCAACCTC 1103
D 8550 CTCCTACTCTCAGCGCA---CAAGGAGCAGCTGGCGCGCTCTCTACTCAGCTCTCTG 8606
QY 1104 CAGCTTAACCTGAGTGGGCGAGGAGCTGTGGAGATCATCTTTCTGGGCTCAAGGCT 1163
D 8607 AGGCGACGCTGACTGGCTGCGGAGCTCGTGGAGACCATCTTTCTGGGTTCCAGGCGC 8666
QY 1164 AGGACATCAGACACACACTCTGAGGACACACCGCTATTCGGGTGATATCTGGCAGATCGG 1223
D 8667 TGGATTCAGGAGCTCCCGCAGGTTCGCCGCTGCGCAGCGCTACGGCAATGCGG 8726
QY 1224 CCCCTGTTCCAAAGCTGCTGTGAACCATGAGAGTGCAGAGTGCCAAATATGTCAAGTCTCAG 1283
D 8727 CCGCTGTTTCTGGAGCTGTGGGAACCAACCGCGAGTGCCCTACGGGTGCTCTCTCAAG 8786
QY 1284 TCACATTGAGGTTTCGAACAGCAACCAACAGGTGACAGATGCTT----- 1331
D 8787 ACGACTGCCCGCTGGAGCTGCGGTTCACCCAGCAGCGCGGTGCTGTGCCCGGAGAAG 8846
QY 1331 -----GAACACAGCCCGCCGAGGAGGAGACAGACCCCGCTGCTGTGTCGAT 1358
D 8847 CCGCAGGGCTCTGTGGCGCGCCCGGAGGAGGAGACAGACCCCGCTGCTGTGTCGAT 8906
QY 1359 TTGCTCCGCTGCACAGCTCCCTGGCAGGTATATGTTTCTTTCGGGCTGCTCTCTG 1418
D 8907 CTGCTCCCGCAGCAGCAGCGCCCTGGCAGGTGTACGGCTTCGTCGGGCTGCTGTCGCG 8966
QY 1419 AAGTGTGTGCTGCTAGTCTCTGGGTACAGGCACAATAGCGCGCTTCTTTAAGAAC 1478
D 8967 CGGTGTGTGCGCCAGGCTCTGGGCTCCAGSCACAACAGCGCGCTTCTCTCAGGAAC 9026
QY 1479 TTAAGAAGTTTCATCTCGTTGGGAAATAGCGCAACTCATCTGAGGAATGATGTGG 1538
D 9027 ACCAAGAGTTTCATCTCCCTGGGGAAGCATGCCAAGCTTCGCTGAGGAGTGTGAGTGG 9086
QY 1539 AAGATGAAGTAGAGGATTCGCTGCTCCGAGCAGCGCGGGGAGGAGCCGCTGTCCCC 1598
D 9087 AAGATGAGCGCTGGGAGTGTGCTGGCTGCGAGGAGCGCCAGGTGAGGAGTGGTGGCC 9146
QY 1599 GCTGAGGAGCAGCTCTGAGGG 1620
D 9147 GTCGAGGGCGCCAGGCGCCAGAG 9168

```

RESULT 8
AF121948 15332 bp DNA PRI 11-APR-1999
LOCUS

DEFINITION Homo sapiens telomerase reverse transcriptase (TERT) gene, partial cds.
ACCESSION AF121948
NID 94580662
VERSION AF121948.1 GI:4580662
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 15332)
AUTHORS Greenberg,R.A., O'Hagan,R.C., Deng,H., Xiao,Q., Hann,S.R., Adams,R.K., Lichtsteiner,S., Chin,L., Morin,G.B. and DePinho,R.A.
TITLE Telomerase reverse transcriptase gene is a direct target of c-Myc but is not functionally equivalent in cellular transformation
JOURNAL Oncogene 18 (5), 1219-1226 (1999)
MEDLINE 99144726
AUTHORS 2 (bases 1 to 15332)
MORIN,G.B., CARLOS,R. and ADAMS,R.R.
JOURNAL Direct Submission
SUBMITTED (21-JAN-1999) Geron Corporation, 230 Constitution Drive, Menlo Park, CA 94025, USA
FEATURES
source
1..15332
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
1..>15332
/gene="TERT"
1..13446
/gene="TERT"
Join(13447..13720,13825..>15178)
/product="telomerase reverse transcriptase"
Join(13502..13720,13825..>15178)
/gene="TERT"
/codon_start=1
/product="telomerase reverse transcriptase"
/protein_id="AAD24464.1"
/db_xref="PID:94580663"
/db_xref="GI:4580663"
translaton="MPRAPRCRAVRSLRSHYREVLPLATFVRRLLPGQWRVLRQDPP
LLDARGGPEAFETTSVRSYLPNTVDALRSGWGLLLRRGAKNVLAFGFA
LVAPSCAYQVGGPLQLGAATOAAPPHPASPRRLRGCCERAWNSHREAGVPLGLPA
PGARRGGSASRLPLPKPRGCAAEPEPTVGGSCWAHPGRTGRSDRGFCVVSPA
RPAEATSLGALSGRHSHPSVGRHAGPPSTSRPPMDTPCPVYATKFLYS
SGKEDLRPSFLSSILRSLTGARRLVETIFLGRPMWPTPRRLPRLPORYWOMRL
FLELLGNHACPCVGLKTHCPLRAAVTPAGVCAREKPGQSVAAPEEDDPRLVQ
LLRQHSPPQVYGFVFRCLRLRLVPLGLMGRHNERRLNKKFISLGHKAKLSLQEL
TWKMSVRCDAWURSP"
BASE COUNT 4502 a 3772 c 3739 g 3319 t
ORIGIN

Query Match 13.4%; Score 468.8; DB 42; Length 15332;
Best Local Similarity 59.4%; Pred. No. 7.6e-115;
Matches 1046; Conservative 0; Mismatches 522; Indels 194; Gaps 7;
Cyt 23 CCGGGCTTACGACATACCGCGCTCTCTGTTGCCCGCGGCTCCCTCTGCTGCG 82
Db 13486 CCGGGCCACCCCGATCGCGCGCTCCCGCTGCGAGCGCGCTCCCTGCTGCG 13545
Cyt 83 CAGCCGATACCGGAGGTGTGGCGCTGGCAACCTTTGTGGCGCGCTGGGGCCCGAGGG 142
Db 13546 CAGCCACTACCGGAGGTGTGGCGCTGGCAACCTTTGTGGCGCGCTGGGGCCCGAGGG 13605
Cyt 143 CAGCGGCTTGTGCAACCGGGGACCCGAAGATCTACCGCACTTTGTTGCCAATGCT 202
Db 13606 CTGGCGGCTGTGTGCAACCGGGGACCCGCGGCTTTCGGCGCGCTGGTGGCGCTGCT 13665
Cyt 203 AGTGTGCATGCTGGGCTCACAGCTTCCAGCTTTCCTTCCAC----- 255
Db 13666 GTGTGCGTGGCTGGGACGACGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 13725

QY 255 ----- 255
Db 13726 COTCCCCGGGTGCGCGCTCCCGCTGGGTTAGGGCGCGCGGGGGAACACGACATGC 13785
QY 255 ----- 255
Db 13786 GGAGAGAGCGCAGCGGCTTCCCGCGAGGTGTCTCTGCTGAAGGAGCTG 278
QY 279 GTGCGCAGGTTGTGAGAGCTCTGCGAGCGCAAGAGAGAAAGTGTGCTGCTTTGGC 13845
Db 13846 GTGCGCAGGTTGTGAGAGCTCTGCGAGCGCGGCGCAAGAGTGTGCTGCTTTGGC 338
QY 339 TTTGAGTGTCTTAAAGAGCGGCGGCTTCCATGCGCTTCACTAGTAGTGTGCTGCT 13905
Db 13906 TTTGAGTGTCTTAAAGAGCGGCGGCTTCCATGCGCTTCACTAGTAGTGTGCTGCT 398
QY 399 AGTACTTGTCCCAACACTGTTATTGAGACCTTGTGCTGCTGCGCAGTGTGCTGCTGCT 13965
Db 13966 AGTACTTGTCCCAACACTGTTATTGAGACCTTGTGCTGCTGCGCAGTGTGCTGCTGCT 458
QY 459 TTGAGCGGAGTGTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 14025
Db 14026 CTGCGCGGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 518
QY 519 CTGCGCGGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 14085
Db 14086 CTGCGCGGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 578
QY 579 ACCACGGATATCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 14145
Db 14146 GCACTACGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 638
QY 639 AATTTCACTAATCTAGTGTCTTAAACAGATCAAGAGAGTAGTGTGCTGCTGCTGCTGCT 14203
Db 14203 ----- 14203
QY 699 AAACCCCTTGGCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 758
Db 14239 GTCCCCCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 14298
QY 759 GTGCTTCACTAAGAGCGCAGATGCTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 14358
Db 14299 GTGCTTCACTAAGAGCGCAGATGCTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 818
QY 819 AGGAGTGTCTTAAACAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 14358
Db 14359 GGCAGGGTCTTGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 878
QY 879 GAGTGTCTTAAACAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 14418
Db 14419 GTGGTGTCTTAAACAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 929
QY 930 CTGAGTGTCTTAAACAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 14478
Db 14479 AGCGGCTTCTTAAACAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 983
QY 984 CTGTACACCGGCTTCTTAAACAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 14538
Db 14539 GCGGCGGCTTCTTAAACAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1043
QY 1044 CTTTACTCCAGGAGTGTGCGGCTTCTTAAACAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 14598
Db 14599 CTCCTTCTTAAACAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1103
QY 1104 CAGCTTAACTGAGTGTGCGGCTTCTTAAACAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 14655
Db 14656 AGGCGGCTTCTTAAACAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1163
QY 1164 AGGACATCAGGAGTGTGCGGCTTCTTAAACAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 14715
Db 14716 TGGATGCCAGGAGTGTGCGGCTTCTTAAACAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 14775

Mon Aug 16 10:21:24 1999

```

QY 1224 CCCCTGTTCCAAAGCTGCTGGTGAACACATGCAGAGTGCCAAATATGTCAGACTCCTCAGG 1283
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 14776 CCCCTGTTCTGGAGCTGCTGGGAACACGCGCAGTGCCTTACGGGGTGCCTCTCAAG 14835
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1284 TCACATTCAGGTTTCGAACAGCAACACAGAGTGACAGATGCCCTT----- 1331
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 14836 ACGCACTGCCCGCTCGAGCTGCGGTACCCACAGCCGCTGTCTGTGCCCGGAGAAG 14895
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1331 -----GAACACAGCCACCGGACCTCATGGAT 1358
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 14896 CCCAGGGCTCTGTGGCGSCCCCGAGGAGGACACAGACACCCCTCGCTGGCTGGTCAG 14955
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1359 TTGCTCCGCTGCACAGAGTCCCTGGCAGGTATATGTTTCTTCGGGCTGTCTCTGTC 1418
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 14956 CTGCTCCGCGACAGCAGCCCTTGGCAGGTGTACGGCTTCGTGCGGCGCTGCTCGC 15015
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1419 AAGTGTGTCTGTCTGTCTGTGGGTACCGAGCACATGAGCGCCGCTCTTTAAGAAC 1478
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 15016 CGGCTGTGCCCCAGGCTCTGGGCTCCAGGCAAGCAAGCGCGCTTCTCTAGGAC 15075
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1479 TTAAGAAGTTCATCTGTTGGGGAATACGCAAGCTATCACTGCAGGAACATGATGG 1538
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 15076 ACCAAGAAGTTCATCTCCCTGGGGAAGCATGCCAAGCTCTGCTGCAGGAGTGCAGTGG 15135
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1539 AAGATGAAGTAGAGGATTGCCACTGGCTCCGAGCAGCCCGGGAAGACCGTGTCCCC 1598
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 15136 AAGATGAGCTGCGGAGTGCCTTGGCTGGCAGGAGGCCAGGAGGAGGTGTGGCC 15195
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1599 GCTGCAGAGCAGCTGTGAGGG 1620
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 15196 GTGAGGCGCCAGGCCCCAGAG 15217
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 9
HSTERT1 HSTERT1 26414 bp DNA PRI 13-MAY-1999
LOCUS Homo sapiens telomerase reverse transcriptase (TERT) gene, exons
DEFINITION 1-6.
ACCESSION AF128893
NID 94808970
VERSION AF128893.1 GI:4808970
KEYWORDS
SEGMENT 1 of 2
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 26414)
AUTHORS Wick,M., Zubov,D. and Hagen,G.
TITLE Genomic organization and promoter characterization of the gene
encoding the human telomerase reverse transcriptase (hTERT)
JOURNAL Gene 232 (1), 97-106 (1999)
MEDLINE 99267414
REFERENCE 2 (bases 1 to 26414)
AUTHORS Wick,M., Zubov,D. and Hagen,G.
TITLE Direct Submission
JOURNAL Submitted (16-FEB-1999) Dept. of Mol. Biol., Bayer AG Leverkusen,
ZF-BTM, Bldg. Q18, Leverkusen D-51368, Germany
FEATURES
source
1. .26414
/organism="Homo sapiens"
/db_xref="taxon:9606"
1. .>11211
/gene="TERT"
promoter
11212. .11492
/gene="TERT"
exon
11212. .1
/number=1
5'UTR
11212. .11273
/gene="TERT"
exon
11597. .12950
/gene="TERT"
number=2
21762
exon
```

```

/number=3
23852. .24032
/gene="TERT"
/number=4
24720. .24899
/gene="TERT"
/number=5
25394. .25549
/gene="TERT"
/number=6
5774 a 6788 c 7653 g 6198 t 1 others
BASE COUNT
ORIGIN
Query Match 13.4%; Score 468.8; DB 42; Length 26414;
Best Local Similarity 59.4%; Pred. No. 8.1e-115; Indels 194; Gaps 7;
Matches 1046; Conservative 0; Mismatches 52;

QY 23 CCGGGCTTGAGCACATGACCGGCTCTCTGTTGCCCGCGGTGCGTCTCTGCTGCG 82
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 11258 CCGGGCCACCCCGGAGTGCCGCGCTCCCGCTCCGAGCGTGCGTCCCTGCTGCG 11317
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 83 CAGCCGATACCGGAGGTGTGGCGCTGGCAACCTTTGTGCGGCGCTGGGCGCCAGG 142
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 11318 CAGCCACTACCGGAGGTGTGGCGCTGGCAACCTTTGTGCGGCGCTGGGCGCCAGG 11377
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 143 CAGCGGCTTGTGCAACCGGCGGACCCGAAAGATCATACCGCACCTTTGTTGCCAATGCCT 202
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 11378 CTGGCGCTGGTGAAGCGCGGGGACCCGCGGCTTTCCGCGCTGTGTGGCCAGTGCCT 11437
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 203 AGTGTGCATGCATGGGGCTCACAGCTCCACTGCGGACCTTTCTTCCAC----- 255
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 11438 GGTGTGCTGCTGGGAGCGACGCGCCGCCCGCGCGCTCTCTCCGCGAGTGGG 11497
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 255 ----- 255
Db 11498 CCTCCCCGGGTGCGCTGCGGTGGGTGTAGGGGCGCGGGGGGAACACGACATGC 11557
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 255 -----CAGGTGTATCTCTCTGAAAGAGCTG 278
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 11558 GGAGAGCAGCAGCGGACTCAGGGCGCTTCCCGCGAGGTGTCTGCTGAAGAGCTG 11617
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 279 GTGGCCAGGTTGTGCAGAGACTCTGCGAGCGCAACAGAGAGAACCTGTGTGCTTTGCG 338
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 11618 GTGGCCGAGTGTGCAGAGCTGTGCGAGGCGCGGCGAAGAACCTGTGTGCTTCGGC 11677
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 339 TTTGAGCTGCTTAACGAGGCGAGGCGGGCTCCCATGGCTTCACTAGTAGCTGCTGCGT 398
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 11678 TTCGCGTGTGAGCGGGCGCGCGGGCGCGCGGGCGCGCGGGCGCTGCTGCTGCTG 458
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 399 AGCTACTTGGCCCAACACTGTTATTGAGACCTGCGTGTGCTGCTGCTGCTGCTGCTGCTG 11797
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 11738 AGCTACTTGGCCCAACACTGTTATTGAGACCTGCGTGTGCTGCTGCTGCTGCTGCTG 518
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 459 TTGAGCGGAGTGGCGGACGACCTGTGCTACCTGCTGGCAGCTGTGCTTCTTATCTT 11857
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 11798 CTGGCGCGGTGGCGGACGACGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 578
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 519 CTGGTGGCCCCAGCTGTGCTACCAAGTGTGTGGGTCTCCCTGTACCAATTTGTGCC 11917
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 11858 CTGGTGGCTCCCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 638
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 579 ACCACGATATCTGGCGCTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 11975
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 11918 GCCACTCAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 698
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 639 AATTCTACTACTTAGTGTCTTTACACAGATCAAGAGAGTAGTGTGCGCCAGGACCG 12010
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 11975 -----GAACGCGCTTGAACCATAGCTGAGGAGGCGGCTGCTGCTGCTGCTGCTG 758
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 699 AAACCCCTGGCTTGGCATCTCGAGGTACAAAGAGGAGCATCTGAGTCTCACCAGTACA 758
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

Db 12011 GTCCCCCTGGGCGCTGCCAGCCCCGGGTGCGAGGAGCGCGGGGGCAGTGCCACGCCGAAAGT 12070

QY 759 GTCCCTTACAGCTAAGAGGCCAGATGCTATCTCTCCCGAGAGTGGAGGAGGACCCAC 818

Db 12071 CTGCGGTTGCCCAAGAGGCCAGCGCTGCGCCCTGAGCGGAGCGGACGCCGTT 12130

QY 819 AGGCAGGTGCTACCAACCCCATCAGGCAAAATCATGGTGCCAAAGTCCCTGCTCGGTCCCC 878

Db 12131 GGGCAGGGGTCTGTGGGCCCCACCCGGCAGGACCGCTGGACCGAGTGACCGGTGGTTTGT 12190

QY 879 GAGGTGCGCTACT-----GCAGAGAAAGATTGCTTTCTTAAAGAAAGTGTCTGAC 929

Db 12191 GTGGTGTCACTGCCACACCGCGCGAAGAACCACTTTTGAGGGTGCCTCTCTGGC 12250

QY 930 CTGAGTCTCTC---TGGGTGGGTGTGCTGTAAACACAGGCCAGCTCCACAT---CTCTG 983

Db 12251 AGGGGCCACTCCCAACCATCGTGGCCGACGACCGAGCGCCCGCCATCCACATCG 12310

QY 984 CTGTCAACACCCGCCAAAATGCTTTCAGCTCAGGCCATTATTGAGACACACATTTTC 1043

Db 12311 CGGCCACCGTCCCTGGGACAGCGCTGTGCCCGCGGTATACGCCGAGACAAGCACTTC 12370

QY 1044 CTTTACTTCCAGGGAGATGCCCCAAGAGCGTCTTAAACCCCTCATTCCTTACTCAGCACTTC 1103

Db 12371 CTCTACTCTTCAGGCGA---CAAGGAGCAGCTGCGGCCCTCTTCTTACTTACTGCTG 12427

QY 1104 CAGCTAACTTGACTGGGCGCAGGAGACTGGTGAGATCAFTTTCTGGGCTCAAGCGCT 1163

Db 12428 AGGCCAGGCTGACTGGCGCTCGGAGGCTCGTGAGACCATCTTTCTGGGTTCCAGGCC 12487

QY 1164 AGGACATCAGGACCACTTGCAGGACACACCGTCTATCGGTCGATCTGAGATGCGGATGGG 1223

Db 12488 TGGATCCAGGAGACTCCCGCAGGTGGCCGCGCTGCCAGCGCTACTTGGCAATGCGG 12547

QY 1224 CCCTGTTCCAAAGCTGTGTGAACCATGCAGTGGCCATATGTCAGACTCTCAGG 1283

Db 12548 CCCTGTTTCTGGAGCTGTTTGGAGACACGCGCAGTGGCCCTTACGGGTGCTCCTCAAG 12607

QY 1284 TCACATTCCAGGTTTGCAGACGAAACCAACAGGTGACAGATGCGCTT----- 1331

Db 12608 AGCGACTGCCGCTGCGAGCTGCGGTCACCCCAGCAGCGGTGTCTGTGCCGGGAGAG 12667

QY 1331 -----GAACACAGCCCCACCGCACTCATGGAT 1358

Db 12668 CCCCAGGGCTCTGTGGGGCCCCCGAGGAGGAGACAGACCCCGCTGCGCTGGTCAG 12727

QY 1359 TTGTCTCCGCTTGCACAGACTCCCTGGCAGTATATGGTTTTCTTGGGCGCTGCTCTGC 1418

Db 12728 CTGTCTCCCGCAGCAGACAGCCCTGGCAGGTGTACGGCTTCGTGCGGGCTGCTCGCG 12787

QY 1419 AAGTGGTGTCTGTAGTCTCTGGGTCACGACACATGAGCGCCGCTTCTTTAAGAAC 1478

Db 12788 CGGTGTGTGCCCCAGGCTCTGGGCTCCAGGCACACGAACGCCGCTCTCTCAGGAAC 12847

QY 1479 TTAAGAAGTTCATCTGTTTGGGAAATACGGCAAGCTATCACCTGCAGGAACATGATGG 1538

Db 12848 ACCAAGAGTTCATCTCCCTGGGAGGATGCCAAGCTCTCGCTGCAGAGCTGACGTGG 12907

QY 1539 AAGATGAAGTAGAGGATTCACACTGGCTCCGACGACCCCGGGAAGACCGGTGTCGCC 1598

Db 12908 AAGATAGCCTGGGGACTGCGCTTGGTTCGCAGGAGCCCGAGAGGTGGTGGCC 12967

QY 1599 GCTGCAGACCGCTGTGAGG 1620

Db 12968 GTCAGGGCCCGAGGCCGACAG 12989

RESULT	10
10	10

AB016767

LOCUS	AB016767	5491 bp	DNA
DEFINITION	Homo sapiens gene for telomere		
ACCESSION	AB016767		
NID	94239869		

VERSION	AB016767.1	GI:4239869	
KEYWORDS	telomerase transcriptase	htERT	
SOURCE	Homo sapiens DNA		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homiinae; Homo.		
AUTHORS	Takakura,M., Kyo,S., Kanaya,T., Takeda,J. and Inoue,M.		
TITLE	1 (bases 1 to 5491)		
JOURNAL	Direct Submission		
REFERENCE	Submitted (04-AUG-1998) to the DDBJ/EMBL/GenBank databases. Masahiro Takakura, Kanazawa University, School of Medicine, Department of Obstetrics and Gynecology; 13-1, Takaramachi, Kanazawa, Ishikawa 920-0934, Japan (E-mail: takakura@med.kanazawa-u.ac.jp, Tel:81-76-265-2425, Fax:81-76-234-4266)		
AUTHORS	Takakura,M., Kyo,S., Kanaya,T., Hirano,H., Takeda,J., Yutsudo,M. and Inoue,M.		
TITLE	Cloning and characterization of human telomerase catarytic subunit (htERT) gene promoter		
JOURNAL	Cancer Res. (1999) in press		
FEATURES	Location/Qualifiers		
source	1..5491		
GC_signal	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	3250..3255		
GC_signal	/note="telomerase transcriptase promoter"		
gene	3331..3336		
	/note="telomerase transcriptase promoter"		
	3338..5491		
5'UTR	/gene="htERT"		
	3338..3414		
exon	/gene="htERT"		
	/evidence=experimental		
	3338..3633		
	/gene="htERT"		
	/number=1		
CDS	Join(3415..3633,3738..>5088)		
	/gene="htERT"		
	/codon_start=1		
	/product="telomerase transcriptase"		
	/protein_id="BAA74724.1"		
	/db_xref="PID:d1038457"		
	/db_xref="PID:g4239870"		
	/db_xref="GI:4239870"		
	/translation="MPRAPRCVRSLLRSHRYEVLPLATFVRLGPGQGNRLVQRGDP AAFRALVAQIVCPVMDPAPSPFQVSCIKELVARVQLCERGAKNVLAFGFA LLDGARGGPFEATTSVRSYLPNTVTDALRGAWGLLRWDDVLVHLLARCALFV LPAPSCAYQGSPLVOLGNAOTAPPPHAGSRRLGGERAWNHVSREAGVPLGLPA RPAREATSLGASLUPLPKRRGAPEPTPVGQSWAHPTGTPSDRGFCVVPSPA FSGKEQGFSLSSLRPLITLGAARLVETIFLGSPPGCGLPAGCPACPSATGKCGPC WSKDQLGTRSPATGCSRRRTARCELSPOQVPSGRSPRALWRPRTQTIPVAMCS KVSATAPLAGVRLRAGLPAPAGAPRLGLQAQRRLNKKAFISLGHAKLSLQELT WKNVSROCAWLRRSP"		
intron	3634..3737		
	/gene="htERT"		
	/number=1		
exon	3738..5088		
	/gene="htERT"		
intron	/number=2		
	5089..>5491		
	/gene="htERT"		
	/number=2		
BASE COUNT	954 a	1729 c	1677 g 1131 t
ORIGIN			

LOCUS	AB016767	5491 bp	DNA	PRI	20-FEB-1999
DEFINITION	Homo sapiens gene for telomerase transcriptase, partial cds.				
ACCESSION	AB016767				

QY 23 C C C G G C C T T G A G C A C A A T G A C C C G G C T C C T G T T G C C C C G G G T C T C T G T G C G 82


```

Query Match      4.4%;   Score 154.8;   DB 11;   Length 4356;
Best Local Similarity 78.2%;   Pred. No. 7.5e-31;
Matches 186;   Conservative 0;   Mismatches 52;   Indels 0;   Gaps 0;

QY  23  CCGGCGCTTGAGCAATGACCCGCTCTCTGTTGCCGCGGTGCGCTCTCTGCTGCG 82
      |||||
Db   3981 CCGGCGCCACCCCGCGATCGCGCGGTCCCGCTCCGAGCGGTGCGCTCTCTGCTGCG 4040
      |||||

QY  83  CAGCCGATACCGGGAGGTGTGGCGGTGCGAACCTTTGTGCGGCGCCTTGGGCGCCGAGGG 142
      |||||
Db   4041 CAGCCACTACCGGAGGTGTGCGGTGCGCCAGCTTCGTGGGCGCCTTGGGCGCCCGAGGG 1400
      |||||

QY  143 CAGCGCGCTTGTGAAACCGGGGACCCGAGATCTACCGCACTTTGTTGTTGCCAATGSCCT 202
      |||||
Db   4101 CTGCGCGCTGGTGACGCGGGGACCCGGCGCTTTTCGCGCGCTGGTGCCAGTGCCT 4160
      |||||

QY  203 AGTGTGCATGCACCTGGGGCTCACAGCTTCCACCTGCGGACCTTTCCCTTCCACAGGTG 260
      |||||
Db   4161 GGTGTGCTGTCCTGGGACGACGCGCGCCCCCGCGCCCTCTCTTCGCGCAGGTG 4218
      |||||

```

RESULT	13
AF098956	
LOCUS	
DEFINITION	AF098956 2043 bp DNA
ACCSSION	Homo sapiens telomerase reverse transcriptase (TERT) gene, promoter
NID	AF098956 PRI
VERSION	94226057
KEYWORDS	
SOURCE	AF098956.1 GI:4226057
ORGANISM	human.

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
REFERENCE
AUTHORS
TITLE
JOURNAL

1 (bases 1 to 2043)
Horikawa, I., Cable, P. L., Afshari, C. and Barrett, J. C.
Cloning and characterization of the promoter region of human
telomerase reverse transcriptase gene
Cancer Res. 59 (4), 826-830 (1999)
99151529

2 (bases 1 to 2043)
Horikawa, I., Cable, P. L., Afshari, C. and Barrett, J. C.
Direct Submission
Submitted (15-OCT-1998) Laboratory of Molecular Carcinogenesis,
National Institute of Environmental Health Sciences, 111 T. W.
Alexander Drive, P.O. Box 12233, Research Triangle Park, NC 27709
USA

```

FEATURES             Location/Qualifiers
     source            1..2043
                        /organism="Homo sapiens"
                        /db_xref="taxon:9606"
                        /chromosome="5"
                        /map="5p15.33"
     gene              1..>2043
                        /gene="TERT"
     promoter          1..1665
                        /gene="TERT"
     mRNA              1666..>1939
                        /gene="TERT"
     CDS               /product="telomerase reverse transcriptase"
                        1721..>1939
                        /gene="TERT"
                        /note="catalytic subunit"

```

Query Match	5.5%	Score 194.6;	DB 12;	Length 201;
Best Local Similarity	98.0%;	Pred. No. 1.2e-41;		
Matches 197;	Conservative 0;	Mismatches 4;	Indels 0;	Gaps 0;
Qy 1686	TTCTTTTACATCACAGAGAGCACATTC	CAGAGACAGGCTCTCTCTACCGTAAAGT	1745	
Db 1	TTCTTTTATGTACGGGAGAGCACATTC	CAGAGACAGGCTCTCTCTACCGTAAAGT	1745	
Qy 1746	GTGTGGACCAAGCTGCAGAGCATTTG	GAGTTCAGGCAACACCTTGAGAGAGTGC	1805	
Db 61	GTGTGGAGCAAGCTGCAGAGCATTTG	GAGTTCAGGCAACACCTTGAGAGAGTGC	1805	
Qy 1806	GAGCTGTCAAGAAGGAGTTCAGGCAT	CACCAAGGACACCTTGCTAGCCATGCCAT	1865	
Db 121	GAGCTGTCAAGAAGGAGTTCAGGCAT	CACCAAGGACACCTTGCTAGCCATGCCAT	1865	
Qy 1866	AGACTCGCTTCAATCCCCAAG	1886		
Db 181	AGACTCGCTTCAATCCCCAAG	201		

RESULT	12	
AF097365		
LOCUS		
DEFINITION		
	AF097365	4356 bp DNA
		Homo sapiens telomerase reverse transcriptase (TERT) gene, promoter and partial cds.
ACCESSION		PRI
NID		03-FEB-1999
VERSION		
KEYWORDS		
SOURCE		
	AF097365.1	GI:4210970
		human.
ORGANISM		

REFERENCE	
AUTHORS	Eutheria; Chordata; Craniata; Vertebrata; Mammalia; Primates; Carnivora; Canidae; Felidae; Mustelidae;
TITLE	Catarrhini; Hominoidea; Homo. 1 (bases 1 to 4356) The human telomerase catalytic subunit hTERT: organization of the gene and characterization of the promoter Hum. Mol. Genet. 8 (1), 137-142 (1999)
JOURNAL MEDLINE REFERENCE	99105927
AUTHORS	2 (bases 1 to 4356) Cong, Y.S., Wen, J. and Bacchetti, S.
TITLE	Direct Submission Submitted (05-OCT-1998) Pathology, McMaster University, 1200 Main St. W., Hamilton, ON L8N 3Z5, Canada
FEATURES	Location/Qualifiers 1
source	

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="5"
1..>4356
/gene="TERT"
1..3996
/promoter
/gene="TERT"
1aln(<3997..4215,4320..>4356)
/gene="TERT"
/product="telomerase reverse transcriptase"
join(3997..4215,4320..>4356)
/gene="TERT"
/codon_start=1

```

AF128893.1:24720..24899,AF128893.1:25394..25549.
3783..3878,4859..4944,7430..7543,9528..9599,11471..11659,
15461..15587,16468..16529,19716..19840,20622..20759,
21295..21399)

```

/codon_start=1
/product="telomerase reverse transcriptase"
/protein_id="RAD12786.1"
/db_xref="PID:94226058"
/db_xref="GI:4226058"
/translation="MPRAPRCRAVRLSLRSHREVLPATFVRLGPGQNRLVQGGDP
AAAFRALVQCLNCVSDA 393 t

```

[illegible]

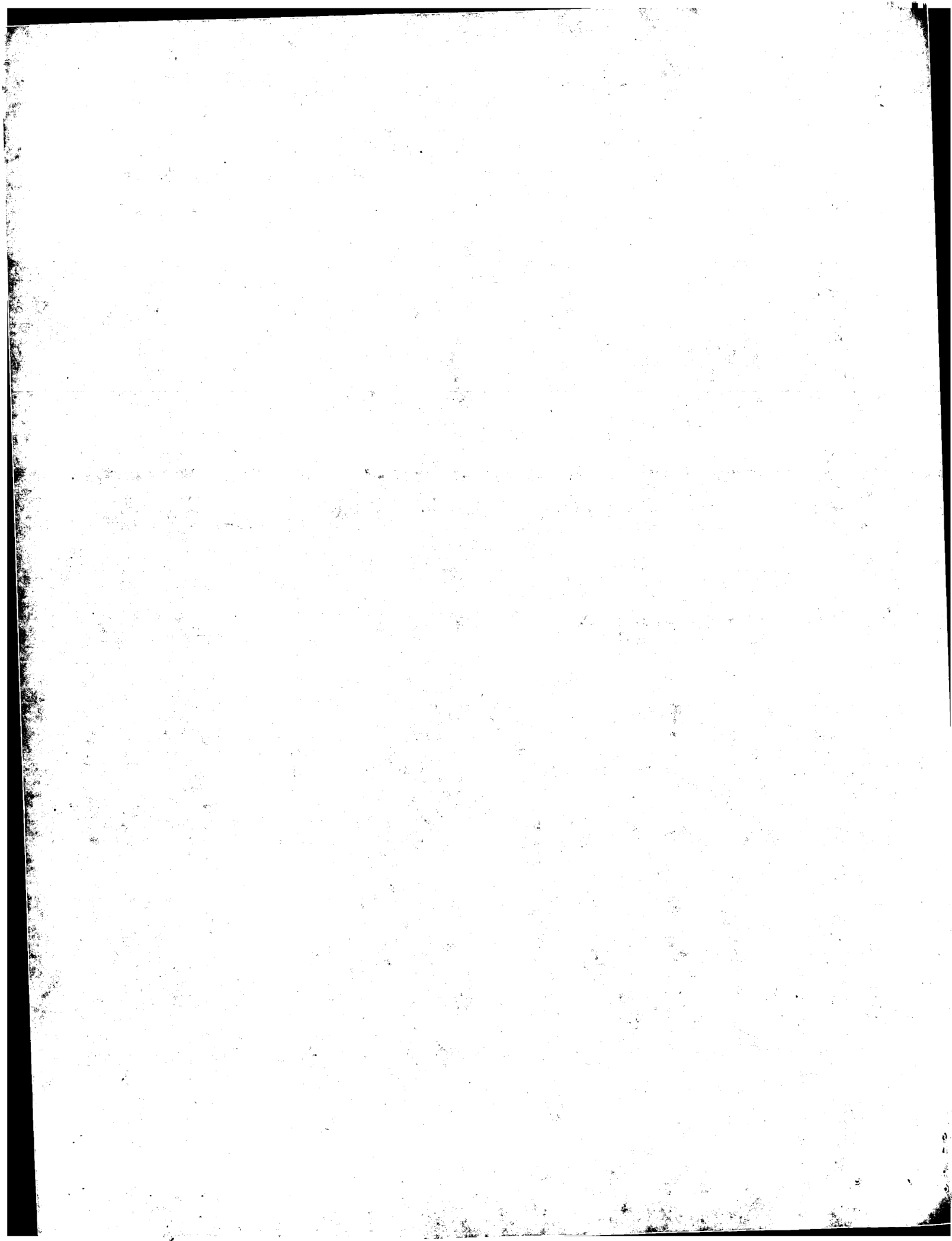
RESULT	14		PRI	13-MAY-1999
HSTERT2		DNA		
LOCUS		25138 bp		
DEFINITION		Homo sapiens telomerase reverse transcriptase (TERT) gene, exons 7-16 and complete cds.		
ACCESSION				
NID		AF128894		
VERSION		94808971		
KEYWORDS		AF128894.1	GI:4808971	
SOURCE		2 of 2		
SEGMENT		human.		
ORGANISM		HOMO SAPIENS		
		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;		
		Eutheria; Primates; Catarrhini; Hominoidea; Homo.		
REFERENCE		1 (bases 1 to 25138)		
AUTHORS		Wick,M., Zubov,D. and Hagen,G.		
TITLE		Genomic organization and promoter characterization of the gene encoding the human telomerase reverse transcriptase (hTERT)		
JOURNAL		Gene 232 (1), 97-106 (1999)		
MEDLINE		99267414		
REFERENCE		2 (bases 1 to 25138)		
AUTHORS		Wick,M., Zubov,D. and Hagen,G.		
TITLE		Direct Submission		
JOURNAL		Submitted (16-FEB-1999) Dept. of Mol. Biol., Bayer AG Leverkusen ZF-BTM, Bldg. Q18, Leverkusen D-51368, Germany		

[illegible]

exon	3783..3878	/gene="TERT"	TTWKSVRDCAWLRSPSGVCPAAEHRRLREEILAKFLHLWMS FLEILGHSHQACQYGVLLKTHCPLRAAVTPAAGCAAREKPOGS SGDEQIURLBGLLUSLSPSLTGARRLVETIFLGRVPMGPGP PSPARRGGSASRSILPLKPKRPRGAAPPEPTPVGQSWAHPGI LLDAGRGSPPEAFTTVSYSLPNTVTDALRSGAWGLLRVLC AAAFRALVACLVCVWDARPPPAAPFQVSCLEKELVARVLQV /translation="MPRAPRCRAVRSLRSHYREVLPATFVFRV /db_xref="GI:4808973"
exon	4859..4944	/gene="TERT"	DDIIRHAKRTFVLVRQADPPPELYFYKVDVTGAYDTIPQD LDDIRHAKRTFVLVRQADPPPELYFYKVDVTGAYDTIPQD VRIYAVQKAAHGHWKAKFSHVIRGKSVQCQGPQGSILST NEASSGLEDFVLFURKCHHVAIRGKSVQCQGPQGSILST IRDGGLLFDVLDLFTVPHLTHAKTFLRTLVRGVPEYICV GGTAFQMPAHGLFPWCGLLDTRTLFVQSDISYATISCI KLFGVLRKCHSLFELDLQVNSLQVCTNIYKILLLOAQRFH FFFLVSDTASLCYSILKAKNAGMSLGNKAGNAPLPSEAVQ YVPLLAGSLPTAQTQLSRKLPDGTTLTALEAAANPALPDSFET /number=7
exon	7430..7543	/gene="TERT"	11471..11659 /number=11 15461..15587 /gene="TERT"
exon	9528..9599	/gene="TERT"	16468..16529 /gene="TERT"
exon	19716..19840	/gene="TERT"	20622..20759 /gene="TERT"
exon	21296..21959	/gene="TERT"	21400..21959 /gene="TERT"
3' UTR			

[illegible]

CDS



GenCore version 4.5
Copyright (c) 1993 - 1998 Compugen Ltd.

OM nucleic - nucleic search, using sw model
Run on: August 13, 1999, 21:51:04 ; Search time 534.86 Seconds
(without alignments)
12893.051 Million cell updates/sec

Title: US-09-042-460-1
Perfect score: 3496
Sequence: 1 GAATTCGGGTGGGAGGCC.....CCGAGCTCGGTACCAAGCTT 3496

Scoring table: IDENTITY_NUC

Searched: 2546578 seqs, 986266752 residues
Database: EST:*

Result No.	Score	Match	Length	DB	ID	Description
1	235.6	6.7	389	31	AA281296	AA281296 zt08g02.r
2	157.8	4.5	409	31	AA311750	AA311750 EST182469
3	91	2.6	375	30	AA200728	AA200728 mu13h09.r
4	91	2.6	503	49	AI645957	AI645957 mu13h09.r
5	57.8	1.7	140	22	R82515	R82515 y19a05.s1
6	56.8	1.6	209	40	AF027363	AF027363 AF027363
7	56.8	1.6	450	49	AF092755	AF092755 AF092755
8	56.2	1.6	503	34	AA514190	AA514190 HFLEST-74
9	56.2	1.6	386	34	AA514191	AA514191 HFLEST-74
10	55.4	1.6	281	49	AF121051	AF121051 AF121051
11	53.6	1.5	418	26	W91597	W91597 MTA.G05.077
12	52.8	1.5	365	49	AF092809	AF092809 AF092809
13	52	1.5	295	41	AF064739	AF064739 AF064739
14	50.6	1.4	266	41	AF064731	AF064731 AF064731
15	50.4	1.4	239	22	U24210	U24210 MMU24210.MO
16	45.2	1.3	555	26	W91680	W91680 MTA.C02.079
17	44.4	1.3	197	49	AJ133823	AJ133823 AJ133823
18	42.8	1.2	474	30	AA262393	AA262393 zsl5b05.r
19	41.2	1.2	469	33	AA443622	AA443622 zsl5b05.r
20	39	1.1	397	33	AA401288	AA401288 zsl5b05.r
21	39	1.1	480	33	AA424462	AA424462 zsl5b05.r
22	39	1.1	377	34	AA504397	AA504397 zsl5b05.r
23	38.2	1.1	399	22	R72912	R72912 y19a05.s1
24	38.2	1.1	425	39	AA856459	AA856459 vx02a08.r
25	38.2	1.1	436	47	AI509561	AI509561 vx02a08.r
26	38	1.1	614	45	U42751	U42751 U42751.Homo
27	37.6	1.1	429	22	R22688	R22688 y08c06.r1
28	37.4	1.1	302	48	AI549853	AI549853 vx02a08.x
29	37	1.1	502	42	AF074090	AF074090 AF074090
30	36.6	1.0	487	41	AA996393	AA996393 S41.HL60
31	36	1.0	470	39	AA859312	AA859312 UI-R-EO-c
32	36	1.0	231	40	AA955313	AA955313 UI-R-EO-c
33	36	1.0	582	43	AI177506	AI177506 EST221138
34	36	1.0	313	43	AI236238	AI236238 EST232800
35	36	1.0	456	46	AA925415	AA925415 UI-R-A1-e
36	36	1.0	558	46	AI408893	AI408893 EST237184
37	36	1.0	422	51	AI705327	AI705327 UI-R-A1-e
38	36	1.0	394	51	AI711279	AI711279 UI-R-A1-e
39	35.8	1.0	278	39	AF037645	AF037645 AF037645
40	35.6	1.0	783	37	AA684336	AA684336 vm89b11.s
41	35.6	1.0	436	44	AI269791	AI269791 qk44d02.x
42	35.6	1.0	246	46	AA926222	AA926222 UI-R-A1-e
43	35.6	1.0	512	50	AU067500	AU067500 AU067500
44	35.4	1.0	349	44	AI298513	AI298513 qm95h02.x
45	35.2	1.0	279	23	H54560	H54560 EST0032.Tes

ALIGNMENTS

RESULT 1
AA281296
LOCUS
DEFINITION zt08g02.r1 NCI_CGAP_GCB1 Homo sapiens cDNA clone IMAGE:712562 5',
ACCESSION AA281296
NID 91924194
VERSION AA281296.1 GI:1924194

14-AUG-1997

us-09-042-460-1.1st

Mon Aug 16 10:21:26 1999

KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
AUTHORS Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
JOURNAL NCI-CGAP Project (CGAP),
COMMENT National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Unpublished (1997)
On Sep 12, 1996 this sequence version replaced gi:1401009.

Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
This clone is available royalty-free through LNL: contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
Insert Length: 2187 Std Error: 0.00
Seq primer: -28m13 rev2 Et from Amersham
High quality sequence stop: 385.

FEATURES
source
1..389
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="NCI CGAP GCB1"
/tissue_type="germinal center B cell"
/lab_host="DH10B"
/note="Vector: p773p-Pac (Pharmacia) with a modified
polylinker. Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
was prepared from human tonsillar cells enriched for
germinal center B cells by flow sorting (CD20+, IgD+),
provided by Dr. Louis M. Staudt (NCI), Dr. David Allman
(NCI) and Dr. Gerald Marti (CBER). cDNA synthesis was
primed with a Not I - oligo(dT) primer
[5'-GTGTTACCACTGAGTGGGCGGCGCTCATTTTTTTTTTTTTTTT-
3']. Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified p773 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 87 a 102 c 123 g 77 t
ORIGIN
Query Match 6.7%; Score 235.6; DB 31; Length 389;
Best Local Similarity 75.6%; Pred. No. 2e-56; Indels 0; Gaps 0;
Matches 292; Conservative
QY 1632 GCTAGCTTCCTGTTCTGGCTGATGGACACATACGTTGGTACAGCTGTAGTCTAGTCTATCTTTT 1691
Db 1 GCCAAGTTCCTGCACTGGCTGATGAGTGTAGCTGCTGAGCTCTCAGTCTCTCTTT 60
QY 1692 TACATCAGAGAGACATTCACGAAGACAGCTCTTCTTACCGTAAGAGTGTCTGG 120
Db 61 TATGTCAGGACACACCTTTCAAGAAGACAGGCTCTTTTCTACCGAAGAGTGTCTGG 1811
QY 1752 AGCAAGCTGAGAGATTGGAGTACAGGCAACCTTGGAGAGTGGCTACCGGAGCTG 180
Db 121 AGCAAGTGCAGCAATTCGATCAGACAGACTTGAAGAGGTTGAGCTGCGGAGCTG 1871
QY 1812 TCACAAGAGAGGTGAGGATCACCAGGACACCTGGCTAGCCATGCCATCTGCACACTG 240
Db 181 TCGGAAGAGAGGTGAGGATCAGGAGCTCGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1931
QY 1872 CGCTTCATCCCAAGCCACAGGCTCGGCGCCATTTGTAACATGAGTTATAGCATGGT 300
Db 241 CGCTTCATCCCAAGCCACAGGCTCGGCGCCATTTGTAACATGAGTTATAGCATGGT 1991
QY 1932 ACCAGAGCTTGGGACAGAGGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 360
Db 301 GCCAAGACGTTCCCGCAGAGAAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 360

QY 1992 TTCAGCATGCTCAACTATGAGCGGAC 2017
Db 361 TTCAGCGTCTCACTACGAGCGGC 386
RESULT 2
AA311750 409 bp mRNA EST 19-APR-1997
LOCUS AA311750
DEFINITION EST182469 Jurkat T-Cells VI Homo sapiens cDNA 5' end, mRNA
sequence.
AA311750
ACCESSION G1964077
NID AA311750.1 GI:1964077
VERSION
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
AUTHORS Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 409)
Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A.,
Bult,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D.,
White,O., Sutton,G., Blake,J.A., Brandon,R.C., Man,Wai,C.,
Clayton,R.A., Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D.,
Fitzgerald,L.M., Fitchugh,W.M., Fritchman,J.L., Geoghegan,N.S.,
Glodek,A., Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S., J.,
Kleek,J.M., Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M.,
Moreno-Palancas,R.F., McDonald,D.A., Nguyen,D.T., Pelligrino,S.M.,
Phillips,C.A., Ryder,S.E., Scott,J.L., Saudek,D.M., Shirley,R.,
Small,K.V., Spriggs,T.A., Utterback,T.R., Weidman,J.F., Li,Y.,
Bednarek,D.P., Caci,L., Cepeda,M.A., Coleman,T.A., Collins,E.J.,
Dimke,D., Feng,D.F., Ferrie,A., Fischer,C., Hastings,G.A.,
He,W.W., Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K.,
Kozak,D.L., Kunsch,C., Hungjun,J., Li,H., Weissner,P., Olsen,H.,
Raymond,L., Wei,Y.F., Wing,J., Xu,C., Yu,G.L., Ruben,S.M.,
Dillon,P.J., Fannon,M.R., Rosen,C.A., Haseltine,W.A., Fields,C.,
Frazer,C.M. and Venter,J.C.
Initial assessment of human gene diversity and expression patterns
based upon 83 million nucleotides of cDNA sequence
Nature 377 (6547 Suppl), 3-174 (1995)
JOURNAL
MEDLINE 96026280
COMMENT On Sep 12, 1996 this sequence version replaced gi:1402064.

Contact: Kerlavage, AR
Bioinformatics
The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850 USA
Tel: 3018699056
Fax: 3018699423
Email: arkerlavet@tigr.org
For clone availability, additional sequence and expression
information related to this EST, please check the TIGR Human Gene
Index (<http://www.tigr.org/tdb/hgi/hgi.html>)
Seq primer: M13 Reverse.

FEATURES
source
1..409
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="Jurkat T-cells VI"
/cell_type="T-lymphocyte"
/note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
XhoI"

BASE COUNT 65 a 120 c 133 g 86 t 5 others
ORIGIN
Query Match 4.5%; Score 157.8; DB 31; Length 409;
Best Local Similarity 75.3%; Pred. No. 3e-34;
Matches 195; Conservative
QY 2611 GTTTTGTGATGACTTCTTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 2670
Db 1 GTTTTGTGATGACTTCTTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 60

Contact: Marra M/Mouse EST Project
WASHU-HHMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.edu
This clone is available royalty-free through LNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:391321
Seq primer: -28M13 rev2 from Amersham
High quality sequence stop: 361.
Location/Qualifiers
1..375
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:639329"
/clone_lib="Soares 2NDMT"
/sex="male"
/tissue_type="Thymus"
/dev_stage="4 weeks"
/lab_host="DH10B"
/note="Vector: pTT73D-Pac (Pharmacia) with a modified
polylinker; Site1: Not I; Site2: Eco RI; lstr strand cdna
was primed with a Not I - oligo(dT) primer [5,
TGTTACCAATCTGAGTGGAGCGCGCGGTTTTTTTTTTTTTTTTTTTT
3']; double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT73 vector. RNA
provided by Dr. Bertrand Jordan. Library went through two
rounds of normalization, and was constructed by Bento

Contact: Maria M/WashU-NCI Mouse EST Project 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu

This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
This read is a RESEQUENCE of a previously sequenced mouse clone
correct orientation)
Seq primer: -40RP from Gibco
High quality sequence stop: 474.

Location/Qualifiers
1. 503
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/map="2"
/clone="IMAGE:639329"
/clone_lib="Soares 2NDMT"
/sex="male"
/tissue_type="Thymus"
/dev_stage="4 weeks"
/lab_host="DH10B"
/note="Vector: pT7R3D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was primed with a Not I - oligo(dT) primer [5'
TGTTACCAATCTGAAGTCGCGCCGCCGGTTTTTTTTTTTTTTTTTTTTT
3']; double-stranded cDNA was ligated to Eco RI adaptors
(pharmacia), digested with Not I and cloned into the Not I

FEATURES
source

FEATURES	SOURCE
High quality sequence stop: 361.	NCBI GenBank
Location/Qualifiers	
1..375	
/organism="Mus musculus"	
/strain="C57BL/6J"	
/db_xref="taxon:10090"	
/	

DEFINITION AF092755 Toxoplasma gondii encystation, subtracted cDNA Toxoplasma gondii cDNA clone CL17, mRNA sequence.

ACCESSION AF092755

NID 94731799

VERSION AF092755.1 GI:4731799

KEYWORDS EST.

SOURCE Toxoplasma gondii.

ORGANISM Toxoplasma gondii.

REFERENCE Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida; Sarcocystidae; Toxoplasma.

AUTHORS Yahaoui,B., Dzlarszinski,F., Bernigaud,A., Slomianny,C., Camus,D. and Tomavo,S.

TITLE Isolation and characterization of a subtractive library enriched for developmentally regulated transcripts expressed during encystation of Toxoplasma gondii

JOURNAL Mol. Biochem. Parasitol. 99, 223-235 (1999)

COMMENT On Mar 10, 1998 this sequence version replaced gi:2948993.

Contact: Stanislas Tomavo
Laboratoire de Chimie Biologique, CNRS UMR 8576
Batiment C9, Universite des Sciences et Technologies de Lille,
59655 Villeneuve d'Ascq, France
Email: Stan.Tomavo@univ-lille.fr.

FEATURES source

Location/Qualifiers

1..450

/organism="Toxoplasma gondii"

/db_xref="taxon:5811"

/clone="CL17"

/clone_lib="Toxoplasma gondii encystation, subtracted cDNA"

/dev_stage="encystation"

/note="Vector: T/A cloning vector pCRII(Invitrogen)"

BASE COUNT 92 a 108 c 142 g 105 t 3 Others

ORIGIN

Query Match 1.6%; Score 56.8; DB 49; Length 450;

Best Local Similarity 96.7%; Pred. NO. 1.8e-05;

Matches 58; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3437 AGGGCAATTCAGCACACTGGCGCGTTACTAGTGGATCGGACCTCGGTACCAAGCTT 3496

Db 65 AAGCCGAATTCAGCACACTGGCGCGTTACTAGTGGATCGGACCTCGGTACCAAGCTT 6

RESULT 8

AA514190/c

LOCUS AA514190 503 bp mRNA EST 14-JUL-1997

DEFINITION HFLEST-741 Human fetal liver (S.Xue) Homo sapiens cDNA, mRNA

sequence.

ACCESSION AA514190

NID 92253714

VERSION AA514190.1 GI:2253714

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS Du,Q., He,F. and Xue,S.

TITLE Human liver ESTs

JOURNAL Unpublished (1997)

COMMENT On May 18, 1995 this sequence version replaced gi:811227.

Contact: Shepu Xue
Dept of Cellular Biology
Institute of Basic Medical sciences/Peking Union Medical College &
Chinese Academy of Medical Sciences(PUMC & CAMS)
5 Dong Dan San Tiao, Beijing, 100005 P.R. China
Tel: 8601-65296459
Fax: 8610-65240529
Email: Xuesp@cdm.imicams.ac.cn

FEATURES source

Location/Qualifiers

1..503

/organism="Homo sapiens"

/db_xref="taxon:9606"

/map="4; 14"

/clone_lib="Human fetal liver (S.Xue)"

/tissue_type="liver"

/dev_stage="fetal"

/lab_host="E.coli DH5a"

/note="Vector: pBluescript SK"

BASE COUNT 128 a 124 c 118 g 133 t

ORIGIN

Query Match 1.6%; Score 56.2; DB 34; Length 503;

Best Local Similarity 95.1%; Pred. NO. 2.9e-05;

Matches 58; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3436 AAGGGCAATTCAGCACACTGGCGCGTTACTAGTGGATCGGACCTCGGTACCAAGCT 3495

Db 67 ATGACCAATTCAGCACACTGGCGCGTTACTAGTGGATCGGACCTCGGTACCAAGCT 8

QY 3496 T 3496

Db 7 T 7

RESULT 9

AA514191

LOCUS AA514191 386 bp mRNA EST 14-JUL-1997

DEFINITION HFLEST-742 Human fetal liver (S.Xue) Homo sapiens cDNA, mRNA

sequence.

ACCESSION AA514191

NID 92253715

VERSION AA514191.1 GI:2253715

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS Du,Q., He,F. and Xue,S.

TITLE Human liver ESTs

JOURNAL Unpublished (1997)

COMMENT On May 18, 1995 this sequence version replaced gi:811228.

Contact: Shepu Xue
Dept of Cellular Biology
Institute of Basic Medical sciences/Peking Union Medical College &
Chinese Academy of Medical Sciences(PUMC & CAMS)
5 Dong Dan San Tiao, Beijing, 100005 P.R. China
Tel: 8601-65296459
Fax: 8610-65240529
Email: Xuesp@cdm.imicams.ac.cn

Seq primer: M13 Forward and Reverse Primer.

Location/Qualifiers

1..386

/organism="Homo sapiens"

/db_xref="taxon:9606"

/map="14"

/clone_lib="Human fetal liver (S.Xue)"

/tissue_type="liver"

/dev_stage="fetal"

/lab_host="E.coli DH5a"

/note="Vector: pBluescript SK"

BASE COUNT 90 a 97 c 100 g 99 t

ORIGIN

Query Match 1.6%; Score 56.2; DB 34; Length 386;

Best Local Similarity 95.1%; Pred. NO. 2.5e-05;

Matches 58; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

us-09-042-460-1.rst

Mon Aug 16 10:21:26 1999

QY 3436 AAGGCGAATTCACACACATGGCGCGGTACTAGTGGATCGAGCTCGGTACCAAGCT 3495
 Db 118 ATGGACGAATTCACACACATGGCGCGGTACTAGTGGATCGAGCTCGGTACCAAGCT 177
 QY 3496 T 3496
 Db 178 T 178
 RESULT 10
 AF121051/c 281 bp mRNA EST 21-APR-1999
 LOCUS AF121051 Homo sapiens liver fetal Homo sapiens cDNA clone ITG-L8,
 DEFINITION mRNA sequence.
 ACCESSION AF121051
 NID 94630719
 VERSION AF121051.1 GI:4630719
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1. (bases 1 to 281)
 AUTHORS Zhu, T., Zhang, L., Liu, Z., and Zhang, J.
 TITLE Homo sapiens chromosome 11p13 EST
 JOURNAL Unpublished (1999)
 COMMENT On Mar 20, 1998 this sequence version replaced gi:2980587.
 CONTACT: Zhu T
 Medical College
 Nankai University
 94 Weijin Rd., Tianjin, 300071, People's Republic of China.
 FEATURES source
 1. 281
 Location/Qualifiers
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /map="3: 11p13"
 /clone="ITG-L8"
 /clone_lib="Homo sapiens liver fetal"
 /tissue_type="liver"
 /dev_stage="fetal"
 BASE COUNT 61 a 65 c 101 g 54 t
 ORIGIN
 Query Match 1.6%; Score 55.4; DB 49; Length 281;
 Best Local Similarity 90.8%; Pred. No. 3.4e-05;
 Matches 59; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
 QY 3432 CATGAAGGCGAATTCACACACATGGCGCGGTACTAGTGGATCGAGCTCGGTACCA 3491
 Db 65 CACAGCGCAATTCACACACATGGCGCGGTACTAGTGGATCGAGCTCGGTACCA 6
 QY 3492 AGCTT 3496
 Db 5 AGCTT 1
 RESULT 11
 W91597 418 bp mRNA EST 09-JUL-1996
 LOCUS MTA.G05.077.A MTA adult mouse thymus library Mus musculus cDNA
 DEFINITION clone MTA.G05.077 3' end similar to CT repeat, mRNA sequence.
 ACCESSION W91597
 NID g1408023
 VERSION W91597.1 GI:1408023
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 418)

AUTHORS Nguyen, C., Rocha, D., Granjeaud, S., Bernard, K., Naquet, P. and Jordan, B.R.
 TITLE Gene expression in different cell types of the mouse thymus
 JOURNAL Unpublished (1996)
 COMMENT On Apr 14, 1993 this sequence version replaced gi:785863.
 CONTACT: Jordan BR
 Genome Structure and Immune Functions
 Centre d'Immunologie INSERM/CNRS
 Case 906, 13288 MARSEILLE Cedex 9, FRANCE
 Tel: 330491269496
 Fax: 330491269430
 Email: jordan@ciml.univ-mrs.fr
 This sequence was determined at Engelhardt Institute of Molecular Biology, Moscow, Russia (V Zakhariev)
 Seq primer: SP6.
 FEATURES Location/Qualifiers
 1. 418
 source /organism="Mus musculus"
 /strain="C57Bl/6"
 /db_xref="taxon:10090"
 /map="17q21"
 /clone="MTA.G05.077"
 /clone_lib="MTA adult mouse thymus library"
 /lab_host="MC1061 p3"
 /note="Vector: pcDNA1; Site_1: NotI; Site_2: EcoRI; The cDNA library was constructed from poly(A)+ RNA of an adult mouse thymus by oligo-dT primed reverse transcription. cDNA was selected on gel for size above 800 nucleotides after second strand synthesis, then directionally cloned into adaptor ligation and restriction enzyme cleavage after the pcDNA1 vector (Not I on polyA side, EcoR I on the 5' side)."
 BASE COUNT 111 a 102 c 85 g 114 t 6 others
 ORIGIN
 Query Match 1.5%; Score 53.6; DB 26; Length 418;
 Best Local Similarity 81.6%; Pred. No. 0.00014;
 Matches 62; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
 QY 3421 CCTAGATGACATGAGGGCGAATTCACACACATGGCGCGGTACTAGTGGATCGCA 3480
 Db 265 CTCAAAAAAGATGACGTTCGAAATTCACACACATGGCGCGGTACTAGTGGATCGCA 324
 QY 3481 GCTCGGTACCAAGCTT 3496
 Db 325 GCTCGGTACCAAGCTT 340
 RESULT 12
 AF092809/c 365 bp mRNA EST 03-MAY-1999
 LOCUS AF092809 Toxoplasma gondii encystation, subtracted cDNA Toxoplasma
 DEFINITION gondii cDNA clone CL146, mRNA sequence.
 ACCESSION AF092809
 NID 94731853
 VERSION AF092809.1 GI:4731853
 KEYWORDS EST.
 SOURCE Toxoplasma gondii.
 ORGANISM Toxoplasma gondii
 Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida;
 Sarcocystidae; Toxoplasma.
 REFERENCE 1 (bases 1 to 365)
 AUTHORS Yakhiaoui, B., Dzierzinski, F., Bernigaud, A., Slomianny, C., Camus, D. and Tomavo, S.
 TITLE Isolation and characterization of a subtractive library enriched for developmentally regulated transcripts expressed during encystation of Toxoplasma gondii
 JOURNAL Mol. Biochem. Parasitol. 99, 223-235 (1999)
 COMMENT On Mar 10, 1998 this sequence version replaced gi:2949046.
 CONTACT: Stanislas Tomavo

Laboratoire de Chimie Biologique, CNRS UMR 8576
Batiment C9, Université des Sciences et Technologies de Lille,
59655 Villeneuve d'Ascq, France
Email: Stan.Tomavo@univ-lille1.fr.

FEATURES
source
1. .365
/organism="Toxoplasma gondii"
/db_xref="taxon:5811"
/map="5"
/clone="CL146"
/clone_lib="Toxoplasma gondii encystation, subtracted
cDNA"
/dev_stage="encystation"
/note="vector: T/A cloning vector pCRII(Invitrogen)"
73 a 91 c 112 g 80 t 9 others
BASE COUNT
ORIGIN
Query Match 1.5%; Score 52.8; DB 49; Length 365;
Best Local Similarity 93.1%; Pred. No. 0.00022;
Matches 54; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 3435 GAAGGGCAATTCAGCACACTGGCGCGCTTACTAGTCGAGCTCGGTACCAA 3492
|||||
Db 65 GAAGGGCAATTCAGCACACTGGCGCGCTTACTAGTCGAGCTCGGTACCAA 8

RESULT 13
AF064739
LOCUS
DEFINITION
AF064739 295 bp mRNA EST 30-OCT-1998
similar to protein tyrosine phosphatase, mRNA sequence.
AF064739
NID
VERSION
KEYWORDS
SOURCE
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core
eudicots; Caryophyllales; Caryophyllaceae;
Dianthus.
1 (bases 1 to 295)
Kim, J.Y., Chung, Y.S. and Shin, J.S.
Isolation of differentially expressed genes during flower
development and pigmentation in carnation
Unpublished (1998)
On Jan 17, 1998 this sequence version replaced gi:2045312.
Contact: Jeong Sheep Shin
Plant Molecular Genetics
Graduate School of Biotechnology, University of Korea
136-701 Anam-dong 5/1 Seoul, Korea
Tel: 00 82 2 3250 3430
Fax: 00 82 2 927 9028
Email: jsshine@kucn.korea.ac.kr.
Location/Qualifiers
1. .295
/organism="Dianthus caryophyllus"
/db_xref="taxon:3570"
/clone="cpi-10"
/clone_lib="Carnation mRNA"
85 a 57 c 49 g 104 t
BASE COUNT
ORIGIN

Query Match 1.5%; Score 52; DB 41; Length 295;
Best Local Similarity 100.0%; Pred. No. 0.00033;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3435 GAAGGGCAATTCAGCACACTGGCGCGCTTACTAGTCGAGCTCGG 3486
|||||
Db 243 GAAGGGCAATTCAGCACACTGGCGCGCTTACTAGTCGAGCTCGG 294

RESULT 14
AF064731
LOCUS
DEFINITION
AF064731 266 bp mRNA EST 30-OCT-1998
similar to catalase, mRNA sequence.
AF064731
NID
VERSION
KEYWORDS
SOURCE
ORGANISM
Dianthus caryophyllus
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core
eudicots; Caryophyllales; Caryophyllaceae;
Dianthus.
1 (bases 1 to 266)
Kim, J.Y., Chung, Y.S. and Shin, J.S.
Isolation of differentially expressed genes during flower
development and pigmentation in carnation
Unpublished (1998)
On Jan 17, 1998 this sequence version replaced gi:2045306.
Contact: Jeong Sheep Shin
Plant Molecular Genetics
Graduate School of Biotechnology, University of Korea
136-701 Anam-dong 5/1 Seoul, Korea
Tel: 00 82 2 3290 3430
Fax: 00 82 2 927 9028
Email: jsshine@kucn.korea.ac.kr.
Location/Qualifiers
1. .266
/organism="Dianthus caryophyllus"
/db_xref="taxon:3570"
/map="2; 2; 2; 15"
/clone="cpi-2"
/clone_lib="Carnation mRNA"
70 a 69 c 58 g 69 t
BASE COUNT
ORIGIN

Query Match 1.4%; Score 50.6; DB 41; Length 266;
Best Local Similarity 93.0%; Pred. No. 0.00077;
Matches 53; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 3440 GCGAATTCAGCACACTGGCGCGCTTACTAGTCGAGCTCGGTACCAAGCTT 3496
|||||
Db 209 GCGAATTCAGCACACTGGCGCGCTTACTAGTCGAGCTCGGTACCAAGCTT 265

RESULT 15
U24210
LOCUS
DEFINITION
U24210 239 bp mRNA EST 03-OCT-1995
MMU24210 Mouse cerebellum Mus musculus cDNA clone PBS114, MRNA
sequence.
U24210
NID
VERSION
KEYWORDS
SOURCE
ORGANISM
Mus musculus
house mouse.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 239)
Wei, J., Hodes, M., Wang, Y., Feng, Y., Ghetti, B. and Dlouhy, S.
Direct cDNA selection from hybridization of mouse cerebellar cDNA
libraries against microdissection DNA pool in MMU16C3-C4 region
Unpublished (1995)
Contact: Wei, J.
Medical Genetics
IUPUI

Mon Aug 16 10:21:26 1999

975 W. Walnut St., Indianapolis, IN 46202, USA

Insert Length: 1421 Std Error: 0.00

High quality sequence stop: 221.

FEATURES

source

Location/Qualifiers

1. .239

/organism="Mus musculus"

/strain="B6CBA"

/db_xref="taxon:10090"

/map="MMU16C3-C4 region; MMU16C3-C4 region"

/clone="PBS114"

/clone_lib="Mouse cerebellum"

/note="Newborn (P1) mice."

66 a 55 c 59 g 59 t

BASE COUNT

ORIGIN

Query Match 1.4%; Score 50.4; DB 22; Length 239;

Best Local Similarity 90.0%; Pred. No. 0.00082;

Matches 54; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 3437 AGGGCGAATTCACACACTGGCGCGTTACTAGTGGATCCGATCCGTTACCAAGCTT 3496

Db 180 AGGGCGAATTCACACACTGGCGCGTTACTAGTGGATCCGATCCGTTACCAAGCTT 239

Search completed: August 13, 1999, 22:05:56

Job time: 892 sec

GenCore version 4.5
Copyright (c) 1993 - 1998 CompuGen Ltd.

QM nucleic - nucleic search, using sw model

Run on: August 13, 1999, 11:20:47 ; Search time 79.57 Seconds
(without alignments)
10992.479 Million cell updates/sec

```

Title: US-09-042-460-1
Perfect score: 3496
Sequence: 1 GAATTCCGGGTGGAGGCC.....CCGAGTCGGTACCAAGTTT 3496

```

Scoring table: IDENTITY_NUC

Searched: 311585 seqs, 125096042 rest.doc

Database : N_Geneseq 36.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query No.	Score	%		Length	DB	ID	Description
			Match					
1	1600.8	45.8	3955	1	V22379		Human telomerase r	
2	1600.8	45.8	4023	1	V60320		Human telomerase g	
3	1600.8	45.8	4042	1	V72117		Human catalytic te	
4	1599.2	45.7	3798	1	V27876		Human telomerase p	
5	1594.4	45.6	4037	1	V23248		Human telomerase r	
6	1323.2	37.8	3855	1	V23232		Human telomerase r	
7	1264.6	36.2	2848	1	V27872		Human telomerase p	
8	992.6	28.4	2357	1	X155923		cdna encoding a ca	
9	826.8	23.6	2089	1	V72136		Human catalytic te	
10	7715	20.5	2175	1	X155925		Nucleotide sequenc	
11	714.4	19.8	2176	1	V232380		Human telomerase r	
12	691.4	18.4	2171	1	V22436		Human telomerase r	
13	647	18.5	3346	1	V60321		Human telomerase r	
14	582.8	16.7	3972	1	V72122		Alternatively spli	
15	468.8	13.4	4335	1	V16979		Human catalytic te	
16	336.2	9.6	949	1	V27875		Human telomerase r	
17	275.4	7.9	523	1	X15926		Human telomerase p	
18	235.6	6.7	395	1	X15924		Nucleotide sequenc	
19	228.8	6.5	1012	1	V72124		Human catalytic te	
20	119.2	3.4	1153	1	V72122		Human catalytic te	
21	80.8	2.3	182	1	V16980		Human telomerase r	
22	60.8	1.7	1935	1	T35333		Peroxisome prolifer	
23	58.6	1.7	1242	1	T89896		Giam cyclin-select	
C 24	58	1.7	9837	1	T96851		Intron 21 of human	
C 25	56.8	1.6	8136	1	T63236		Plasmid pCR(TW)3la	
C 26	56.8	1.6	6115	1	T63235		Eukaryotic express	
27	56.8	1.6	1272	1	V19124		Nucleotide sequenc	
C 28	55.4	1.6	5120	1	V01426		Plasmid pCMV-SmgSt	
C 29	55	1.6	3984	1	T07310		Vector plasmid pTO	
C 30	55	1.6	2178	1	T15929		Neural cell adhesi	
C 31	55	1.6	9592	1	T37652		Adenovirus vector	
C 32	55	1.6	5540	1	T62548		19Sv-hpOMCdeIACHT	
C 33	55	1.6	10443	1	T71319		Plasmid pTB72 enco	
C 34	55	1.6	5446	1	X39297		Plasmid pCDNA3. pl	
C 35	55	1.6	6420	1	V15625		Sequence of pGEM C	
36	54.8	1.6	219	1	V48407		Human Flt4 ligand	
37	54.6	1.6	908	1	V68843		DNA molecule encod	
38	54.4	1.6	404	1	T69381		Murine metastatic	
39	54	1.5	94	1	Q89696		Vector back-end ol	
40	54	1.5	6216	1	V01435		Plasmid pRC/ASK8-S	
41	53.4	1.5	573	1	T70422		Human rhodopsin cd	
42	53.4	1.5	1982	1	V10366		Wheat low affinity	
43	53.4	1.5	4525	1	V69746		Nucleotide sequenc	

Eukaryotic express
Eucaryotic (CMV) e

ALIGNMENTS

RESULT 1

V22379	ID	V22379 standard; cDNA; 3955 BP.
V22379	AC	V22379;
DT	DT	13-AUG-1998 (first entry)
DE	DE	Human telomerase reverse transcriptase encoding cDNA.
KW	KW	Human; telomerase reverse transcriptase; hTERT; diagnosis;
KW	KW	prognosis; cell proliferation; cancer; ageing; ribonucleoprotein; ss.
OS	OS	Homo sapiens.
FT	FT	Key
FT	FT	CD5
FT	FT	Location/Qualifiers
FT	FT	5C
FT	FT	5C

```
Location/Qualifiers
56. .3454
/*tag= a
/product= "telomera
```

GB2317891-A.
/product="telomerase reverse transcriptase"

08-APR-1998.
01-OCT-1997; 020890.
14-AUG-1997; US-915503.
01-OCT-1996; US-724643.
18-APR-1997; US-844419.
25-APR-1997; US-846017.
05-MAY-1997; US-851843.
03-MAY-1997; US-854050.
14-AUG-1997; US-911312.
14-AUG-1997; US-912951.
(GERO-) GERON CORP.
(UYTE-) UNIV TECHNOLOGY CORP.
Andrews WH, Cech TR, Chapman KB, Morin GB, Nakamura T, Harley CB; WPI: 98-171633/16.
P-PSDB; W46957.
Pure and recombinant human Telomerase Reverse Transcriptase and its variants - are useful in the diagnosis, prognosis and treatment of cell proliferation conditions especially cancer and ageing
Claim 5; Fig 16; 387pp; English.
The present sequence encodes human telomerase reverse transcriptase (hTERT), which is a ribonucleoprotein. The present invention also describes the following methods: (A) determining whether a test compound is a modulator of hTERT, by detecting the change in hTERT recombinant protein or polynucleotide, on administration of the compound; (B) preparation of recombinant telomerase by contacting a protein preparation of hTERT with a telomerase RNA component; (C) detection of the hTERT RNA or protein in a sample by binding a relevant probe to the sample and detecting the complex formed or in the case of RNA detection, amplifying the product and correlating the presence of complex or amplification product with presence of hTERT in the sample; or increasing the proliferation of a vertebrate cell by increasing hTERT expression; and (E) the use of an agent that causes an increase in cell proliferation to create a medicament that inhibits ageing. A protein preparation of hTERT and the polynucleotide encoding hTERT can be used in the manufacture of medicaments for inhibiting the effect of ageing or cancer. Inhibitors of telomerase activity can be used to treat conditions that are associated with high telomerase activity. A protein preparation of hTERT can also be used in the new methods.

Sequence	3955 BP;	648 A;	1353 C;	1251 G;	703 T;
----------	----------	--------	---------	---------	--------

```
Query Match          45.8%;      Score 1600.8;      DB 1;      Length 3955;
Best Local Similarity 68.9%;      Pred. No. 0;
Matches 2379; Conservative 0; Mismatches 957;
```

9
QY 23 CCCGSCCTTGAGCACAATGACCGCGTCTCTCGTTGCCCGGGTGCCTCTCTGTCCG 82
||||||| | ||| ||||| |||||
db 40 CCCGGCACACC CGCGATGCGCGCGCTCCCGCTGCCGACCGGTGCTCCCTTCGCG 99

D	b	1150	TCTGGGTTCCAGGCCCTGGATGCCAGGACTCCCGCAGAGTGTGGCCGCCTGCCCCACGG	1209
Q	y	1208	ATACTGGCAGATGCGGCCCTGTTTCCAACAGCTGCTGGTGAACCATCAGAGTGCCCAATA	1267
D	b	1210	CYATGGCAAATGCGGCCCTGTTTCTGGAGTCGCTTGGAAACCACGCGAGTGCCCTTA	1269
Q	y	1268	TGTCAGACTCTTCAGGTTCACATTGCAAGTTTTGGAACACAAACAAGGTGACAGATGC	1327
D	b	1270	CGGGTGCTCTCAAGACGACTGCCGCTGGAGCTGGGTGACCCACGAGCGGGTGT	1329
Q	y	1328	CT-----TGAAACACAGGCC	1342
D	b	1330	CTGTGCCGGGAGAACGCCAGGGCTCTGTGGCGGCCCGCAGGAGGAGGACACAGACCC	1389
Q	y	1343	ACGCACTCATGGATTGCTCCGCTGCACAGCACTGCTCCGCGAGGTATATGTTTTCT	1402
D	b	1390	CCGTGCGCTGGTGACGTGCTCCGCCAGCACAGACCCCCTGGCAGGTGTACGGCTTGT	1449
Q	y	1403	TCGGGCTGCTCTGCAAGGTGGTGTGCTAGTCTCTGGGTACCAGSCACAATGAGCG	1462
D	b	1450	CGGGCTGCTGCGCGCGGTGGTGGCCCGAGGCTCTGGGGTCCAGGCACAACGAACG	1509
Q	y	1463	CCGCTCTTTAAGAATTAAGAAGTTCACTCGTTGGGAAAATACGGAAGTATFACCT	1522
D	b	1510	CCGCTTCTCAGGAACACCAAGAAGTTCACTCCCTGGGGAACATGCCAAGCTCTCGCT	1569
Q	y	1523	GCAGCACTGATGTGGAAGATGAAGTAGAGATGTCACCTGGCTCCGAGACGCCGGG	1582
D	b	1570	GCAAGAGTGCCTGGAAGATGACGCTCGGGACTGCGCTGGCGAGGAGCCGAGG	1629
Q	y	1583	GAAGACCGCTGCCCGCTGCAGAGCACCGCTGCAGGAGAGGATCCTGGCTAGTTCTCT	1642
D	b	1630	GGTGGCTGTGTCGCGCGCGAGACCGCTCTCGGTGAGGAGATCTTGCCAAAGTTCTCT	1689
Q	y	1643	GTCTGGCTGATGGACACATACGTGATGAGTGCCTTAGTTCATCTTTTACATCACAGA	1702
D	b	1690	GCACTGGCTGATGAGTGTAGCTGCTGAGCTGCTAGGCTCTTTTATGTACGGA	1749
Q	y	1703	GAGCACTTCCAGAGAACAGCGCTCTTCTACCGTAAGAGTGTGGACAAAGCTGCA	1762
D	b	1750	GACCGTTTCAAAGAACAGGCTCTTTTCTACCGGAAGAGTGTCTGGACAAAGTTGCA	1809
Q	y	1763	GAGCATTGGAGTCAGCAACACCTTGAGAGGTGCGGTACGGGAGCTGTCCACAGAGA	1822
D	b	1810	AAGCATTGGAAATCAGACAGCTTTGAAGAGGTGAGCTGCGGGAGCTGTGGAAGAGA	1869
Q	y	1823	GGTCAGCATCACAGAGAACCTGGCTAGCCATGCCATCTCAGACTGCGCTTCATCCC	1882
D	b	1870	GGTCAGCAGCATCGGGAAGCCAGGCGCCGCTGTGAGCTCCAGACTCCGCTTCATCCC	1929
Q	y	1883	CAAGCCCAAGCGCTGGGGCCATTGTGAACATGAGTTATAGCATGGGTACAGAGCTTT	1942
D	b	1930	CAAGCTTGAGGGGTGGCGCGATGTGAACATGGACTAGCTGCTGGGAGCCAGACGTT	1989
Q	y	1943	GGGAGAGGAAGACAGGCCCGCAGCATTTACCCAGCTCTCAACACTCTCTTACAGCATGT	2002
D	b	1990	CCGAGAGAAAAGAGGCGGAGCGCTCACCTCGAGGGTGAAGGCACTGTTTACGCTGCT	2049
Q	y	2003	CAACTATAGCGGACAAAAACATCTCACTTATGGGTCTTCTGTACTGGGTATGAATGA	2062
D	b	2050	CAACTACAGAGGGGGCGGGCCCGCCCGCTCTGGCGCTCTGTGTGGGCTGGACGA	2109
Q	y	2063	CATCTACAGGACCTGGCGGGCTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2122
D	b	2110	TATCCAGGGGCTGGCGACACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2169
Q	y	2123	GATGTACTTTGTAGGCAGATGTGACCGGGGCTATGATGCCATCCCGCAGGTAAAGT	2182
D	b	2170	GCTGTACTTTGTCAAGGTGGATGTGACGGGCGGTACGACACCATCCCGCAGGACGCT	2229
Q	y	2183	GGTGGAGGTGTTGGCAATATGATCAGGCACCTCGGAGAGCAGGTACTATCCCGCAGTA	2242
D	b	2230	CACGGAGGTTCATCGCCAGCATCATCA---AACCCCAACACGATGCTGCTGCTGCTG	2286

Mon Aug 16 10:21:24 1999

Db 283 CTGCTGAAGAGCTGGTGGCCGAGTGTGTCAGAGGCTGTGCGAGCGGGCGGAGAA 342
 QY 323 CGTGTGGCTTTTGGCTTTTGGCTTTTAAAGAGCCAGAGCGGGGCTTCCCATGGCTT 382
 Db 343 CGTGTGGCTTTGGCTTTTGGCTTTTAAAGAGCCAGAGCGGGGCTTCCCATGGCTT 402
 QY 383 CACTAGTAGGCTGCTAGTACTTGGCCCAACACTGTTATATAGACCTTGGCTGCTAGTGG 442
 Db 403 CACCACAGCGTGGCAGCTTACCTGCCAACACAGCTGACCGACACTGCGGGGAGCGG 462
 QY 443 TGCATGATGCTACTGTTAGCGAGTGGGCGACGACTGCTGGTCTACTGCTGGCACA 502
 Db 463 GCGTGGGGGCTGCTGTTGGCGCGCTGAGCGGCGGCGGCGGCGGCGGCGGCT 522
 QY 503 CTGTCCTTTTATCTCTGCTGCGCCCGGAGCTGTGCTACCAAGTGTGGGTCTCCCT 562
 Db 523 CTGCGCGCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 582
 QY 563 GTACCAATTTTGGCCACAGGATCTGCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCT 622
 Db 583 GTACAGCTCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 642
 QY 623 CCGACCGTGGCAGGAATTTCACTAAGCTTAGTCTTAAACAGATCAAGAGCAGTAG 682
 Db 643 GCGTCTGGGATGCT-----GAACGGCGCTTGAACCATAG 675
 QY 683 TCGCCAGGAAGCACCAGGCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 742
 Db 676 CGTCAGGAGCGCGGGTCCCGCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 735
 QY 743 TCTACCAAGTACAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 802
 Db 736 CAGTCCAGCGGAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 795
 QY 803 GGAGGAGGAGCAGCAGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 862
 Db 796 GGAGGAGGAGCAGCAGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 855
 QY 863 TCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 913
 Db 856 TGACCGTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 915
 QY 914 AGGAAAGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 970
 Db 916 GGTGCGCTCTGTCAGCAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 975
 QY 971 CTCACAT---CTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1027
 Db 976 CCGCCCATCCATGCGGCGCCACAGTCCCTGGGACAGCGCTTGTCCCGGTGCTAGCG 1035
 QY 1028 TGAGACAGACATTTCTTTACTTCCAGGAGGAGATGGCCAGAGGCTTAAACCCCTCAT 1087
 Db 1036 CGAGACCAAGCACTTCTCTACTCTCTCAGCGCA---CAAGGAGCAGCTGCGGCCCTCTT 1092
 QY 1088 CCTACTCAGCAACTCCAGCTTAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1147
 Db 1093 CCTACTCAGCTCTTGAAGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1152
 QY 1148 TCTGGGTCAAGGCTTAGGACATCAGGACCACTCTGCGAGACACACCGTCTATCGCGTGG 1207
 Db 1153 TCTGGGTCCAGGCGCTTGGATGCGAGGAGTCCCGCAGGCTTCCCGGCTGCGCCAGCG 1212
 QY 1208 ATACTGGCAGATGGGCGCTTGTCCACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1267
 Db 1213 CTACTGGCAATGGGCGCTTGTCTGAGGCTGCTTGGAGACCAAGCGCAGTGGCCCTTA 1272
 QY 1268 TGTACAGCTCTCAGGTACATGCTGAGGTTTTCGAACAGCAAAACCAAGGTGACAGTGC 1327
 Db 1273 CGGGTGTCTCAGAGCGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1332
 QY 1328 CT-----TGAAACACAGCGCC 1342
 Db 1333 CTGTGCGCGGAGAGCCCGCAGGCTGTGTGGCGCGCGCGGAGGAGGACAGACAGCC 1392

QY 1343 ACCGACCTCATGATTTGCTCCGCTGTCAGCAGAGTCCCTGGCAGGTATATGTTTCT 1402
 Db 1393 CCGTCCCTGGTGCAGTGTCCGCGCAGCAGACAGCCCTGGCAGGTATATGTTTCT 1452
 QY 1403 TCGGGCTCTCTCTGCAAGTGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1462
 Db 1453 GCGGGCTGCTGCTGCGCGCTGGTGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1512
 QY 1463 CCGCTCTTTAAAGACTTAAAGAACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1522
 Db 1513 CCGCTCTCTCAGAAACACCAAGAGTTCTCTCCCTGGGAAAGCATGCAAGCTCTCCT 1572
 QY 1523 GCAGGAAGTGTGCAAGATGAAAGTAGAGTGGCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1582
 Db 1573 GCAGGAGTGTGCAAGATGAAAGTAGAGTGGCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1632
 QY 1583 GAAGGAGCTGTCCCGCTGCGAGCAGCCTGCTGAGGAGAGATCTGCTGCTGCTGCTGCTGCTGCTGCT 1642
 Db 1633 GGTGGCTGTGCTGCGCGCAGCAGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1692
 QY 1643 GTTCTGCTGATGAGACATACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1702
 Db 1693 GCATGCTGATGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1752
 QY 1703 GAGCAGCTCCAGAGACAGGCTTCTTCTTACCGTAAAGTGTGCTGCTGCTGCTGCTGCTGCTGCT 1762
 Db 1753 GACCAAGTCTTCAAGAGACAGGCTTCTTCTTACCGTAAAGTGTGCTGCTGCTGCTGCTGCTGCT 1812
 QY 1763 GAGCAGCTCCAGAGACAGGCTTCTTCTTACCGTAAAGTGTGCTGCTGCTGCTGCTGCTGCTGCT 1822
 Db 1813 GAGCAGCTCCAGAGACAGGCTTCTTCTTACCGTAAAGTGTGCTGCTGCTGCTGCTGCTGCTGCT 1872
 QY 1823 GGTAGGATCAGCAGGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1882
 Db 1873 GGTAGGATCAGCAGGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1932
 QY 1883 CAAGCCCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1942
 Db 1933 CAAGCCCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1992
 QY 1943 GGGCAGAGGAGCAGGCGGCGGCTTCCAGGCTTCAAGACTCTCTTTCAGCATGCT 2002
 Db 1993 CCGCAGAGAGAGGCGGCGGCTTCCAGGCTTCAAGACTCTCTTTCAGCATGCT 2052
 QY 2003 CAACTATGAGCGAGCAAAACATCTCACTATGAGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2062
 Db 2053 CAACTATGAGCGAGCAAAACATCTCACTATGAGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2112
 QY 2063 CATCTACAGACCTGCGGCGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2122
 Db 2113 TATCCACAGGCGCTGCGCAGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2172
 QY 2123 GATGCTTGTAAAGCAGATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2182
 Db 2173 GCTGCTTGTCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2232
 QY 2183 GGTGAGGCTTGTGCAATATGATCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2242
 Db 2233 CACGAGGCTCATGCGCAGCATCTCA---AACCCAGAGACAGTCTGCTGCTGCTGCTGCTGCTGCT 2289
 QY 2243 TGCAGTGTCCGAGAGATAGCCAGGCGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAG 2302
 Db 2290 TGCAGTGTCCGAGAGATAGCCAGGCGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAG 2349
 QY 2303 CACCTCTCTGACCTCCAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAG 2362
 Db 2350 TACCTTGACAGACCTCCAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAG 2403
 QY 2363 TGCAGTGTCCGAGAGATAGCCAGGCGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAG 2422
 Db 2404 GACCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCT 2463

Mon Aug 16 10:21:24 1999

Db	407	CACCACGCGTGGCAGCTAACCTGCCAACACGCTGACCGACGCACTCGCGGGGAGCGG	466
Qy	443	TGCATGATGCTACTGTGTAGCCGAGTGGCGACGACCTGCTGGTACTACTCTGCTGCACA	502
Db	467	GGCGTGGGGGTGCTGTCTGGCGCGGTGGCGACACGTCGTGGTTCACTGCTGGCAG	526
Qy	503	CTGTGCTTTTATCTTCTGTGTGCCCGCCAGCTGTGCTACCAAGTGTGGTTCGCCCT	562
Db	527	CTGGCGCTCTTTGTGTGTGGTCCCGAGTGGCTGCGCTACCGAGTGTGGGCGCGCGCT	586
Qy	563	GTACCAATTTGTGCCACACGGATATCTGGCCCTCTGTGTCGCTAGTATACAGGCCAC	622
Db	587	GTACCACTCGCGCTGCCACTCAGGCGCGCGCCCGCCACACGCTAGTGGACCCCGAG	646
Qy	623	CGACCCGTGGCAGGAATTTCACTAACCTTAGTCTTATACAACAGATCAAGACGATAG	682
Db	647	GGGTCTGGATGC-----GAAGGCGCTTGAACCACTAG	679
Qy	683	TGCGCAGGAAGCACCGAAACCCCTGGCCTTGCCATCTCGAGGTACAAAGGCACTCTGAG	742
Db	680	CGTCAGGAGGCGCGGGTCCCGCTGGGCTGCCAGCCCCGGGTGCGAGGAGGCGGGGG	739
Qy	743	TCCTACCACTACAAGTGTGCTTCAGCTTAAGAAAGCCAGATGCTATCTTCCCGAGAGT	802
Db	740	CAGTGCACCGCAAGTCTGCGGTGCCCAAGAGCCCGAGGCGTGGCGCTGCCCTGAGCC	799
Qy	803	GGAGGAGGAGCCCCACAGGAGTGTCTACCAACCCCATCAGGCAAAATCATGGTGGCCAG	862
Db	800	GGAGGAGGAGCCGCTGGCGAGGGTCTCTGGGCCCCACCGGGCGAGGACGCGTGACCGAG	859
Qy	863	TCCTGCTCGGTCGCCCGAGGTGCCTACT-----GCAGAGAAGATTTGTCTTTCTAA	913
Db	860	TGACCGTGGTTTCTGTGTGGTGTCACTTCCAGTCCAGACCGCGGAAGACCACTCTTTGGA	919
Qy	914	AGGAAAGGTGTCTGACCTGAGTCTCTC---TGGGTGGTGTGCTGTAAACACAGCCAG	970
Db	920	GGGTGCGCTCTCTGGCACGCGCACTCCCAACCATCCGTGGCGGCCACGACACCGCGG	979
Qy	971	CTCCACAT--CTCTGCTGCACACCGCGCAAAATGCCTTTCAGCTTCAGGCCATTTAT	1027
Db	980	CCCCCATCCACATCGCGCCACAGTCCCTGGACAGCCCTTGTCCC CGGTGTACGC	1039
Qy	1028	TGAGCAGACATTTCTTTACTTCCAGGGGAGTGGCCAAAGAGCGCTTAACCCCTCAT	1087
Db	1040	CGAGCCAAAGCACTTCTCTACTCTCTCAGGCGA---CAAGGAGACGCTCGCGCCCTCT	1096
Qy	1088	CTTACTCAGCAACTCCAGCTTACTTGTGTGGGCGCAGGACACTGGTGGAGATCATCTT	1147
Db	1097	CCCTACTCAGTCTCTGAGGCCACGCTGACTGCGCTCGGAGGCTCGTGGAGACCATCTT	1156
Qy	1148	TCTGGGCTCAAGGCGCTAGGACATCAGGACCATCTTCGAGGACACACCCGCTTATCGGCTG	1207
Db	1157	TCTGGTTCCAGGCGCTTGATGTCAGGGAATCCCGCAGGTTGCCCCGCTCGCCACGG	1216
Qy	1208	ATACTGGCAGATGGCGCCCTGTTCACACAGCTGCTGGTGAACCATGACAGTGGCAATA	1267
Db	1217	CTACTGGCAATATGGGCCCCCTGTTTCTGAGACTGCTTGGAAACCAACGCGCACTGCCCTA	1276
Qy	1268	TGTCAGACTCTTCAGGTACATTCGAGGTTTCGAACAGCAAAACAACAGGTGACAGATGC	1327
Db	1277	CGGGTGTCTCTCAAGAGCACTGCCCGCTGCGAGCTGCGGTTCACCCACAGCAGCGGTGT	1336
Qy	1328	CT-----TGAAACACCAAGCC	1342
Db	1337	CTGTGCCCGGAGAACCCAGGCGTCTGTGGCGCCCCCGAGGAGGAGCACAGACCC	1396
Qy	1343	ACCGCACTCATGATTTGCTCCGCCTGCACAGAGTCCCTTGGCAGGTATATGGTTTTCT	1402
Db	1397	CGTGGCTGGTGCAGTGTCTCCGACGACACAGCAGCGCCCTTGGCAGGTGTACGGCTCGT	1456
Qy	1403	TCGGGCTGTCTCTCCAGGTGGTGTCTGTAGTCTCTGGGGTACCAGGCACATGAGCG	1462
Db	1457	GGCGGCTGTCTTGCCTCGCGGTGGTGGCCCCAGGCTCTGGGGCTCCAGGCACACGACG	1516

QY	1463	CGCCTCTTTAAGRACATTAAAGAAGTTTCATCTCGTTCGGGAATATACGGCAAGCTATCACT	1522
Db	1517	CGCGTTCTCTCAGGAACACCAAGAAGTTCATCTCCCTGGGAAGCATGCCAAGCTCTCGCT	1576
QY	1523	CGAGAACTGATGTGGAAAGATGAAGTAGAGGATGCCACTGGCTCCGCAGCAGCCCGGG	1582
Db	1577	CGAGGAGCTGACGTGGAAGATGAGCGTTCGGGAGTCCGCTTGGCTGGCAGAGGCCACGG	1636
QY	1583	GAAGACCGTGTCCCGCTGCAGAGACACCGCTCTGAGGGAGAGGATCTCGCTACGTTTCCCT	1642
Db	1637	GGTTGGCTGTGTTCGGCGCGCAGACACCGTCTCGCTGAGGAGATCTCTGGCCAAGTTTCT	1696
QY	1643	GTTCGGCTGATGGACATACGTTGGTACAGCTGCTTAGTTCATCTTTTACATCAGAGA	1702
Db	1697	GCACCTGGCTGATGAGTGTGTAGCTGCTCGAGCTGCTCAGGCTTTTCTTTTATGTCACGGA	1756
QY	1703	GAGCACATTCAGAGAACAGAGCTCTTCTTCTACCGTAAAGTGTGTGGAGCAAGCTGCA	1762
Db	1757	GACACGTTTCAAAAGAACAGGCTCTTTTCTACCGAAGAGTGTCTGGAGCAAGTTGCA	1816
QY	1763	GAGCATTTGGAGTACGCCAACACTTGAAGATGGCGTACGGGAGCTGTCCACRAGAGGA	1822
Db	1817	AAGCATTTGGAATCAGACACGACCTTGAAGAGGTGTCAGCTGCGGGAGCTGTTCGGAAGCAGA	1876
QY	1823	GGTCAGGCATCACCGAGACACCTGGCTAGCCATGCCCATTCTGCAGACTTCGGTTCATCCC	1882
Db	1877	GGTCAGGCACCATCGGAAGCCAGGCCGCCCTGCTGACGTCACAGACTCCGCTTCATCCC	1936
QY	1883	CAAGCCCAACGGGCTCGGGGCCATTGTGAACATGAGTTATAGCATGGGTACCAGAGCTTT	1942
Db	1937	CAAGCCTTGAGGGGCTCGGGCGGATTGTGAACATGAGACTACGTCGTGGGAGCCAGAACGTT	1996
QY	1943	GGGCAGAGAAGCAGGCCACGATTTTCACCCAGCGCTCTCAAGACTCTCTTCAGCATGCT	2002
Db	1997	CCGCAGAGAAAAGGCGCCAGCGCTCTCACTTCGAGGGTGAAGGCACCTGTTCAAGCTGCT	2056
QY	2003	CAACTATGAGCGGCAAAACATCCTCACTTATGGGCTTCTCTGTACTGGGTATGAATCA	2062
Db	2057	CAACTAGAGCGGGCGGGCGCCCGGCCCTCTGGCGCCCTCTGTGCTGGGCTGGACGA	2116
QY	2063	CATCTACAGGACCTCGGGGCGCTTTGTGCTGCGTGTGCGTGTCTGGACAGACACCCAG	2122
Db	2117	TATCCAGAGGCTTGGCGCACCTTCGTTGCTGCGTGTGCGGGCCAGGACCGCGCGCTGA	2176
QY	2123	GATGTACTTTGTTAAGCAGATGTGACGGGGCCCTATGATGCCATCCCCAGAGGTAAAGCT	2182
Db	2177	GCTGTACTTTGTCAGGTGGATGTGAGGGGCGGTAGCACACCATCCCCCAGGACAGGCT	2236
QY	2183	GGTGGAGGTTGTGGCAATATGATCAGGCACCTCGGAGAGCAGCTPACTGTATTCGCGCAGTA	2242
Db	2237	CAGGAGGTATCGCCAGCATCA---AACCCCAAGAACGCTACTTGGCTGGGTCTGGTA	2293
QY	2243	TGCATGTGTCGGAGAGATAGCAAGCCAGTCCACAAGTCCCTTTTAGGAGACAGGTCCAC	2302
Db	2294	TGCGTGTGTCAGAGAGGCGGCCATGAGGACAGCTCCGAGGCTTCAAGAGCCACGCTC	2353
QY	2303	CACCTCTCTGACCTCCAGCCATACATGGGCGAGTTCTTAAAGCATCTGCAGGATTCAGA	2362
Db	2354	TACCTTGACAGACCTCCAGCGGTACATCGCAGATTTCTGGCTCACCTGCAG-----GA	2407
QY	2363	TGCGAGTGCACTGAGGAACCTCGTTGTATCGAGCAGAGCATCTCTATGAATGAGAGCAG	2422
Db	2408	GAGCACCCGCTGAGGAGTATGCGCTGCTCATGAGCAGAGCTCTCTCCCTGATGAGGCCAG	2467
QY	2423	CAGCAGCCTGTTTGACTTCTTCCCTGCATCTCTGCTCACAGTGTCTGTAAGATTTGTTGA	2482
Db	2468	CAGTGGCCTTCGACGCTCTTCTCAAGCTTCTATGTGCCACACCGCGTGGCATCAGGGG	2527
QY	2483	CAGGTGCTATPCCAGTGGCCAGGCATCCCCCAGGGCTCCAGGCCTATCACCCCTGCTCTG	2547
Db	2528	CAAGTCTCTAGTCCAGTGGCAGGGGATCCCGCAGGGCTCCATCTCTCCAGCTCTGCTG	2587

Mon Aug 16 10:21:24 1999

QY 503 CTGTGCTCTTTATCTTCTGTCGCCGCCAGCTGTGCTTACCAGTGTGTGGTCTTCCCT 562
 DB 532 CTGCGGCTCTTTGCTGCTGCTCCAGCTGCGCTTACCAGTGTGTGGGCGCGCGCT 591
 QY 563 GTACCAAAATTTGGCCACACAGGATATCTGGGCTGTGTGCTGCTTACAGGCCAC 622
 DB 592 GTACCAAGCTCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 651
 QY 623 CCAGCCGCTGGCAGGAATTTCACTAACTTAGTCTTACACAGATCAAGAGCAGTAG 682
 DB 652 GCCTGTGGGATGCT-----GACGGGCTGGAACCATAG 684
 QY 683 TCGCCAGGAGACCCGAAACCCCTGGCTTGGCATCTCGAGGTACAAAGAGCATCTGAG 742
 DB 685 CGTCAGGAGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 744
 QY 743 TCTCAGCAGTACAGTGTGCTTCACTAAGAGGCGAGATGCTATCTTCTCCGAGAGT 802
 DB 745 CAGTGCAGCGGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 804
 QY 803 GGAGGAGGACCCACAGGAGGTGTACCAACCCCATCAGGCAATCATGGTGCCAG 862
 DB 805 GGACGCGACCGCTTGGCGAGGCTCTGCGGCCACCGCGGAGGAGCGCTGGACCGAG 864
 QY 863 TCCTGCTCGGTCCTCCCGAGTGCCTACT-----GCAGAGAAAGATTGTCTTTCTAA 913
 DB 865 TGACCGTGGTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 924
 QY 914 AGGAAAGGTGCTGACCTGAGTCTCTC-----TGGGTGCTGCTGCTGCTGCTGCT 970
 DB 925 GGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 984
 QY 971 CTCCACAT-----CTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1027
 DB 985 CCCCCCATCCACATGCGGCGCACACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1044
 QY 1028 TGAGACAGACATTTCTTTTCTTCTCAGGAGATGGCCAGAGCGCTGTAAACCCCTCAT 1087
 DB 1045 CGAGACCAAGACATCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1101
 QY 1088 CTTACTACGACACTTCCAGCTTAACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1147
 DB 1102 CTTACTGCTCTCTGAGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1161
 QY 1148 TCTGGCTCAAGCTAGGACATCAGGACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1207
 DB 1162 TCTGGCTTCCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1221
 QY 1208 ATACTGGCAGATCGGCGCTGCTTCCACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1267
 DB 1222 CTACTGGCAATGCGGCGCTGCTTCTGAGCTGCTTGGCAACCGGCTGCTGCTGCTGCTGCT 1281
 QY 1268 TGTCACTCTCTCAGTGCATGAGGTTTCCAAACAGCAACCAACAGGCTGACAGATGC 1327
 DB 1282 CGGGGTGCTCTCAAGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1341
 QY 1328 CT-----TGACACACAGCC 1342
 DB 1342 CTGTGCGCGGAGAACCCAGGCTCTGTGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1401
 QY 1343 ACGCAGCTCATGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1402
 DB 1402 CCGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1461
 QY 1403 TCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1462
 DB 1462 CGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1521
 QY 1463 CCGCTTCTTAAAGACTTAAAGACTTAAAGACTTAAAGACTTAAAGACTTAAAGACTTAAAGACT 1522
 DB 1522 CCGCTTCTCAGGAACACCAAGAGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1581
 QY 1523 GCAGGAATGATGTGGAAGATGAAAGTAGAGATTGCCACTGCTGCTGCTGCTGCTGCTGCTGCT 1582

DB 1582 GCAGGAGCTGACGTGGAACATGAGCGTGGGACTGCGCTTGGCTGCGCAGGAGCCAGG 1641
 QY 1583 GAAGGACCGTGTCCCGCTGCGAGACCGCTGAGGAGAGGATCTTGGCTTACCTTCTTCT 1642
 DB 1642 GGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1701
 QY 1643 GTTCTGCTGATGACACACATAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1702
 DB 1702 GCCTGCTGATGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1761
 QY 1703 GAGCACATCCAGAGAACAGGCTCTTCTTACCTGAGAGTGTGTGAGCAAGCTGCA 1762
 DB 1762 GACCACTTTCAAAAGAACAGGCTCTTCTTACCTGAGAGTGTGTGAGCAAGTGTGCA 1821
 QY 1763 GAGCATTTGAGTTCAGGCAACACCTTGAAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCT 1822
 DB 1822 AAGCATTTGATATGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1881
 QY 1823 GGTGAGGATCACCAGACACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1882
 DB 1882 GGTGAGGATCACCAGACACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1941
 QY 1883 CAAGCCCAAGGCTGCGGCGCTTGTGAACATGAGTGTGCTGCTGCTGCTGCTGCTGCTGCT 1942
 DB 1942 CAAGCTGACGGCTGCGGCGCTTGTGAACATGAGTGTGCTGCTGCTGCTGCTGCTGCTGCT 2001
 QY 1943 GGGCAGAGGAGGAGGCGGCGCTTGTGAACATGAGTGTGCTGCTGCTGCTGCTGCTGCTGCT 2002
 DB 2002 CCGCAGAGAAAGAGGCGGCGCTTGTGAACATGAGTGTGCTGCTGCTGCTGCTGCTGCTGCT 2061
 QY 2003 CAACTATGAGCGGAGGCGGCGCTTGTGAACATGAGTGTGCTGCTGCTGCTGCTGCTGCTGCT 2062
 DB 2062 CAACATGAGCGGCGGCGGCGCTTGTGAACATGAGTGTGCTGCTGCTGCTGCTGCTGCTGCT 2121
 QY 2063 CATCTACAGGAGGCGGCGGCGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2122
 DB 2122 TATCCACAGGCGGCGGCGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2181
 QY 2123 GATGTACTTTTGAAGCAGATGACCGGCGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2182
 DB 2182 GCTGTACTTTTGAAGCAGATGACCGGCGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2241
 QY 2183 GGTGAGGTTGTTGCAATATGATCAGGCACTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2242
 DB 2242 CACGAGGTCATCGCCAGCATCATCA---AACCCAGAACACGCTACTGCTGCTGCTGCTGCT 2298
 QY 2243 TGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2302
 DB 2299 TGCCGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2358
 QY 2303 CACCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2362
 DB 2359 TACTTGAACAGCTTCCAGCGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2412
 QY 2363 TGCCAGTGCATGAGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2422
 DB 2413 GACCAAGCGCTGAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2472
 QY 2423 CAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2482
 DB 2473 CAGTGGCTTCTGAGCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2532
 QY 2483 CAGTGTCTATGAGCAGTGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2542
 DB 2533 CAGTGTCTATGAGCAGTGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2592
 QY 2543 CAGTGTCTATGAGCAGTGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2602
 DB 2593 CAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2652
 QY 2603 GCTTTAGCTTTTGTGATGACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2662

Db 2653 GCTCCTGCGCTTGGTGATGATTTCTTGTGTGACACCTCACCCTACCCACGCGAAAC 2712
QY 2663 CTTCTCAGACACCTGGTCCATGGGTTCTCTGAGTATGGTGCATGATAAACTTGCAGAA 2722
Db 2713 CTTCTCAGACACCTGGTCCGAGGTGTCCTCTGAGTATGGTGGTGTGAACTTGCAGAA 2772
QY 2723 GACAGTGGTGAATCTCCCTGTGGAGCCCTGGTACCCTGGTGGTGGTGCAGTCCATCAGCT 2782
Db 2773 GACAGTGGTGAATCTCCCTGTAGAGAGAGAGCCCTGGTGGTGGTGGTGGTGGTGGT 2832
QY 2783 GCTCCTCAGTCCCTGTTTCCCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 2842
Db 2833 GCGGCGCCACGCGCTATTCCTCCTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 2892
QY 2843 GTTCTGTGACTACTCAGTGTATGCCAGACCTCAATTAAGACGAGCTCAGCTTCCAGAG 2902
Db 2893 SCAGAGCAGCTACTCCAGCTATGCCGAGCTCCATCAGAGCCAGTCTCAGCTTCAACCG 2952
QY 2903 TGTCTTCAAGCTGGGAAGACATCGGGAACAAGCTCTGTGCGTCTTGGGTTGAAGTG 2962
Db 2953 CGGCTTCAAGCTGGGAGGACATCGCTGCGAACTCTTTGGGCTCTTGGGCTGAAGTG 3012
QY 2963 TCAGGCTTATTTAGACTTGCAGTGAACAGCCTCCAGACAGTCTGCATCAATATATA 3022
Db 3013 TCAGGCTTATTTAGACTTGCAGTGAACAGCCTCCAGACAGTCTGCATCAATATATA 3072
QY 3023 CAAGATCTTCTGCTTCAAGGCTACAGTTCATGATGATGATGATGATGATGATGATGAT 3082
Db 3073 CAAGATCTTCTGCTGAGGCTACAGTTCATGATGATGATGATGATGATGATGATGAT 3132
QY 3083 CCAGCTGTTAGGAAGACCTCAATTTCTTGGGCTATCTCCAGCAGCTTCCCTTTGA 3142
Db 3133 TCAGCAAGTTTGAAGAACCCCAATTTTCTGCGGTATCTCTGACAGCGCTCCCT 3192
QY 3143 CTGCTATGCTATCTGAAGTCAAGATCAGATCAGAAATGACACTAAAG-----GCCTC 3193
Db 3193 CTGCTACTCATCTGAAACCAAGAACGAGGATGTCGCTGGGCGCAAGGCGCGC 3252
QY 3194 TGGCTCTTCTCCTGAAAGCCGACATTTGGCTCTGCTACAGGCTTCTGCTCAAGT 3253
Db 3253 CGGCTCTGCTCCTGAGGCGCTGAGTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3312
QY 3254 GCTGCTCTATCTGCTATCAATATGCTCTCTGGGACCTCTGAGGACGCGCAAAACT 3313
Db 3313 GACTCAGACCGTGTCACTAGTGCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3372
QY 3314 GCTGTGCGGAAGCTCCAGAGCGGACAAATGACCTTAACTGAGTGGAGTGGAGTGGAG 3373
Db 3373 GCTGAGTGGGAAGCTCCCGGGGAGGAGCTGACTGCTTGGAGGCGGAGCAACCGGC 3432
QY 3374 CTAAGCAGAGCTTTCAGACCAATTTGGACTAA 3407
Db 3433 ACTGCCCTCAGACTTCAAGACCATCTCTGACTGA 3466

RESULT 5
V22428

ID V22428 standard; cdna; 4037 BP.

AC V22428.

DT 13-AUG-1998 (first entry)

DE Human telomerase reverse transcriptase encoding cdna refined sequence.

KW Human; telomerase reverse transcriptase; hTERT; TRF; diagnosis;

OS prognosis; cell proliferation; cancer; ageing; ribonucleoprotein; ss.

QW Homo sapiens.

FT Key

FT Location/Qualifiers

FT CDS

FT 56..3454

FT /*tag= a

FT /product= "telomerase reverse transcriptase"

FT /note= "refined sequence"

FN GB2317891-A.

PD 08-APR-1998.

PF 01-OCT-1997; 020890.

PR 14-AUG-1997; US-915503.

PR 01-OCT-1996; US-724643.
PR 18-APR-1997; US-844419.
PR 25-APR-1997; US-846017.
PR 05-MAY-1997; US-851843.
PR 09-MAY-1997; US-854050.
PR 14-AUG-1997; US-911312.
PR 14-AUG-1997; US-912951.
PA (GERO-) GERON CORP.
PA (UYTE-) UNIV TECHNOLOGY CORP.
PI Andrews WH, Cech TR, Chapman KB, Harley C, Lingner J,
PI Morin GB, Nakamura T, Harley CB;
DR WPI: 98-171633/16.
DR P-PSDB: W56113.
PT Pure and recombinant human Telomerase Reverse Transcriptase and its
PT variants - are useful in the diagnosis, prognosis and treatment of
PT cell proliferation conditions especially cancer and ageing
PS Example 1; Fig 74; 387pp; English.
CC The present sequence encodes human telomerase reverse transcriptase
CC (hTERT), which is a ribonucleoprotein. The present invention also
CC describes the following methods: (A) determining whether a test compound
CC is a modulator of hTERT, by detecting the change in hTERT recombinant
CC protein or polynucleotide, on administration of the compound; (B)
CC preparation of recombinant telomerase by contacting a protein
CC the hTERT RNA or protein in a sample by binding a relevant probe to the
CC sample and detecting the complex formed or in the case of RNA detection,
CC amplifying the product with presence of hTERT in the sample; and (D)
CC amplification product of a vertebrate cell by increasing hTERT
CC expression; and (E) the use of an agent that causes an increase in cell
CC vertebrate cell proliferation to create a medicament that inhibits
CC ageing. A protein preparation of hTERT and the polynucleotide encoding
CC hTERT can be used in the manufacture of medicaments for inhibiting the
CC effect of ageing or cancer. Inhibitors of telomerase activity can be
CC used to treat conditions that are associated with high telomerase
CC activity. A protein preparation of hTERT can also be used in the new
CC methods.
SO Sequence 4037 BP; 682 A; 1362 C; 1275 G; 714 T;

Query Match

Best Local Similarity 45.8%; Score 1594.4; DB 1; Length 4037;

Matches 2375; Conservative 0; Mismatches 971; Indels 108; Gaps 9;

QY 23 CCAGCGCTTACGACCAATGACCCGCGCTCTCGTCCCGCGGCGCTCTCTGTCTGTGG 82
Db 40 CCGGGCCACCCCGCGATGCGCGCGCTCCCGCTCCGAGCGGTGCTGCTGTGG 99
QY 83 CAGCCGATACCGGAGGTGTGGCCGCTGGCAACCTTTGTGCGGCGCTGGGCGCGAGG 142
Db 100 CAGCCCACTACCGGAGGTGTGGCCGCTGGCCAGTTCGTGCGGCGCTGGGCGCGAGG 159
QY 143 CAGCGCGCTGTGTCAACCGGAGCCGCGAGATCTACCCGACTTTGGTGGCCATGGCT 202
Db 160 CTGCGCGCTGTGTGACCGGCGGAGCCCGCGGCTTTCCGCGCGCTGGTGGCGAGTCT 219
QY 203 AGTGTGATGCTACCTGGGCTTCAGCGCTCCAGCTTCGCGACCTTCTCTCCACAGGTCTC 262
Db 220 GGTGTGCTGCTGCGTGGGAGCGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 279
QY 263 ATCCCTGAAGAGCTGTGGCCAGGTTGTGAGAGACTTGTGAGAGCTTGTGAGAGAGAA 322
Db 280 CTGCTTGAAGAGCTGTGGCCGCGAGTGTGAGAGGCTGTGAGAGGCTGTGAGAGGCG 339
QY 323 CGTGTGCTTTTGTGCTTGTAGCTCTTAACGAGCGGAGCGGCGCTTCCCATGGCTT 382
Db 340 CGTGTGCTTGTGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 399
QY 383 CACTAGTAGCTGCTAGTACTTGTCCCAACACTGTTATTGAGACCTCTGCTGTGAGTGG 442
Db 400 CACCAACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 459
QY 443 TGCATGGATGCTACTGTTGAGCCGAGTGGGCGAGCAGCTGCTGCTGCTGCTGCTGCT 502

Db	1510	CGCCTTCTCAGGAACACCAAGAGTTTCTCTCCCTGGGGAAGCATGCCAAGCTCTCGCT	1569
Qy	1523	GCAGGAACTGATGTGAAGATGAAGTAGAGATTGCCACTGGCTCCGAGAGCCCGGG	1582
Db	1570	GCAGGAGCTGACGTGGAAGATGAGCGTGGGACTGCGCTTGGCTGGCAGGAGCCAGG	1629
Qy	1583	GAAGGACCGTGTCCCGCTGAGAGACCGCTGAGGAGAGATCCCTGGCTACGTTCTCT	1642
Db	1630	GGTGGCTGTGTCCCGCGCAGAGCAGCGTCTGCTGAGAGATCTGGCCAAAGTCTCT	1689
Qy	1643	GTTCGTGCTGATGAGACACATACGTGTGAGAGTGTAGTCTATTCTTTATCATCAGA	1702
Db	1690	GCATGGCTGATGAGTGTGATGCTGAGCTGCTGAGCTCTTCTTTTATGTCAGGA	1749
Qy	1703	GAGCACATTCCAGAAACAGAGCTCTTCTTACCCTAGAGTGTGTGGAGCAAGTGA	1762
Db	1750	GACCAGCTTCAAAAGAACAGGCTCTTTTACCAGCCGAGTGTCTGGAGCAAGTTGA	1809
Qy	1763	GAGCATTGGAGTCAAGCAACACCTTGAGAGAGTGGGCTACGGAGCTGTCAAGAGGA	1822
Db	1810	AAGCATTGGAATCAGACAGCATTGAAGAGGGTGCAGCTCGGAGACTGTGCGAAGCA	1869
Qy	1823	GGTCAGGCATCACCAGGACACCTGGCTAGCATGCCATCTGCAGACTGCGCTTATCCC	1882
Db	1870	GGTCAGGACGATCGGNAAGCCAGCCGCCCTGCTGACGTCCAGACTCGCTTATCCC	1929
Qy	1883	CAAGCCCAACCGGCTGGGCGCATTTGAACTGATGATGATGATGATGATGATGATG	1942
Db	1930	CAAGCCTCACGGGCTGGGCGCATTTGAACTGATGATGATGATGATGATGATG	1989
Qy	1943	GGGAGAGAGGAGGAGGCGGCGCATTTTCAAGCAGCTCTCAAGACTCTCTTACGATGCT	2002
Db	1990	CGGAGAGAAAGAGGCGGCGCTCTCACCTCGAGGGTGAAGCACTGTTTACGCGTCT	2049
Qy	2003	CAACTATGAGCGGACAAACATCTTACCTTATGGGCTTCTTACTGGTATGAATGA	2062
Db	2050	CAACTACGAGCGGCGGCGGCTCTTGGGCGCTCTTGGGCGCTTGGAGCA	2109
Qy	2063	CATCTACAGGACTCGGCGGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2122
Db	2110	TATCCACAGGCGCTGGCGACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2169
Qy	2123	GATGTACTTTTGAAGGAGATGTACCGGCGCTTATGATGCCATCCCGAGGTAAGCT	2182
Db	2170	GCTGTACTTTTGAAGGATGTACCGGCGCTTACGACACCATCCCGGAGGACAGCT	2229
Qy	2183	GGTGGAGTGTGCAATATGATGAGGACCTCGGAGAGCAGTACTGTATCGGCACTA	2242
Db	2230	CAGGAGGCTATCGCCAGCATCATCA---AACCCAGAACACGTACTGCTGCTCGGTA	2286
Qy	2243	TGCAGTGGTCCGAGAGATAGCCAAAGGCCAAGTCCCAAGTCTTTAGGAGACAGTAC	2302
Db	2287	TGCGGTGTCAGAAAGCGCGCATGGGACGTCGCAAGGCTTCAAGAGCAGCTCTC	2346
Qy	2303	CACCTCTCTGACTCCAGCCATACATGGCGAGTTCCTTAAAGCATCTCGAGATTCAG	2362
Db	2347	TACCTTGACAGCTCCAGCGCTACATGACAGTGTGCTGGCTACCTCCAG---GA	2400
Qy	2363	TGCCAGTGCATGAGAACTCCCTTGTCTATCGAGAGAGCATCTTATGAATGAGAGAG	2422
Db	2401	GACCAGCGGCTGAGGATGCGCTGCTATCGAGAGAGCTCTCTCTCTGAATGAGGCG	2460
Qy	2423	CAGCAGCGCTGTTGACTTCTCTGCACTTCTGCTGCTGCTGCTGCTGCTGCTGCTG	2482
Db	2461	CAGTGGCTCTTCCAGCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2520
Qy	2483	CAGTGTCTATACGAGTCCAGGCGCATCCCGAGGCTCCAGGCTTCCAGCTTGTCTG	2542
Db	2521	CACTCTCTGCTGAGTCCAGTCCAGGAGTCCCGAGGCTTCCATCTCTCCAGCTGCTG	2580
Qy	2543	CAGTCTGTGTTTCCGAGACATGAGAACAGCTGTTTGTGCTGAGTGTGAGGATGGGT	2602
Db	2581	CAGCTGTCTACGCGACATGGAGAACAGCTGTTTGGCGGAGTTCGGCGGAGCGGCT	2640

Db	460	GGCGTGGGGCTCTGCTGCGCGGTGGGCGACAGCTGCTGTTTCACTCTGCTGGCAG	519
Qy	503	CTGTGCTTTTATCTTCTGTTGGTCCCGAGCTGTGCTTACAGAGTGTGTGGTCTTCCCT	562
Db	520	CTGCGGCTCTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	579
Qy	563	GTACCAAAATTTGGCCACACAGATATCTGGGCGCTCTGTGCTGCTGCTGCTGCTG	622
Db	580	GTACAGCTCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	639
Qy	623	CCGACCGTGGGAGGAATTTACTAATCTAGTGTCTTACACAGATCAAGACAGTAG	682
Db	640	CGCTCTGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	672
Qy	683	TCGCCAGGAGCAGCAACCCCTGGCTTGGCATCTCGAGGTACAAAGAGGATCTGAG	742
Db	673	CGTCAGGAGGCGGGTCCCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	732
Qy	743	TCTCAGCAGTACAGTGTGCTTACAGTGAAGAGGCGAGTGTATCTTGTCCGAGAGT	802
Db	733	CAGTGCAGCGAAGTCTGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	792
Qy	803	GGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	862
Db	793	GGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	852
Qy	863	TCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	913
Db	853	TGACCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	912
Qy	914	AGGAAGGTGCTGACCTGAGTCTCTC---TGGGCTGCTGCTGCTGCTGCTGCTG	970
Db	913	GGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	972
Qy	971	CTCCACAT---CTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1027
Db	973	CCCCCATCCACATCGCGGCGGACAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTG	1032
Qy	1028	TGAGACAGACATTTCTTTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1087
Db	1033	CGAGACCAAGCACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT	1089
Qy	1088	CCTACTCAGCAACCTCCAGCTTAACTGACTGGGCGGAGAGAGTGTGGAGATCATCT	1147
Db	1090	CCTACTCAGCTCTCTGAGCGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1149
Qy	1148	CTGCTGCTCAGGCTTAGACATCAGGACCTCTGCTGCTGCTGCTGCTGCTGCTG	1207
Db	1150	CTGCTGCTCAGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1209
Qy	1208	ATACTGGCAGATGCGGCGCTGCTTCAACAGCTGCTGCTGCTGCTGCTGCTGCTG	1267
Db	1210	CTACTGGCAATGCGGCGCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1269
Qy	1268	TGTCAGACTCTCAGTGTACATGAGTGTGCAACAGCAACACAGGAGTACAGATGC	1327
Db	1270	CGGGGTGCTCTCAAGACGACTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1329
Qy	1328	CT-----TGAAACACAGGCC	1342
Db	1330	CTGTGCGGAGAGGCGGCTCTGTGGGCGGCGGAGGAGGAGGAGGAGGAGGAGG	1389
Qy	1343	ACCGACCTCATGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1402
Db	1390	CGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1449
Qy	1403	TCGGGCTGCTGCTGCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1462
Db	1450	CGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1509
Qy	1463	CCGCTTCTTAAAGAACTTAAAGAGTTTCTCTGTTGGGGAATACGGCAAGCTACT	1522

263	ATCCCTGAAGAGCTGGTGGCCAGGGTTGTCAGAGACTCTGCAGCGCAACGAGAGAA	322	Y
264			
265			
266			
267			
268			
269			
270			
271	CTCCCTGAAGGAGCTGGTGGCCGAGTCTCGAAGGCTGTGCGAGCGCGCGAAGAA	350	b
272			
273			
274			
275			
276			
277			
278			
279			
280			
281			
282	CGTGCTGGCTTTTGGCTTTGAGCTGCTTAAAGGCGCAGAGCGGGCTCCATGCGCTT	382	Y
283			
284			
285			
286			
287			
288			
289			
290			
291	CGTGCTGGCTTCGGCTTCGCGCTGCTGAGCGGGCGCGCGGGGCCCGCCGAGGCTT	410	b
292			
293			
294			
295			
296			
297			
298			
299			
300			
301			
302			
303			
304			
305			
306			
307			
308			
309			
310			
311			
312			
313			
314			
315			
316			
317			
318			
319			
320			
321			
322			
323			
324			
325			
326			
327			
328			
329			
330			
331			
332			
333			
334			
335			
336			
337			
338			
339			
340			
341			
342			
343			
344			
345			
346			
347			
348			
349			
350			
351			
352			
353			
354			
355			
356			
357			
358			
359			
360			
361			
362			
363			
364			
365			
366			
367			
368			
369			
370			
371			
372			
373			
374			
375			
376			
377			
378			
379			
380			
381			
382			
383			
384			
385			
386			
387			
388			
389			
390			
391			
392			
393			
394			
395			
396			
397			
398			
399			

Db	1341	CTGTGCGGGGAGAACCCCCAGGGCTCTGTGGCGCCCCGAGGAGGAGACACAGACC	1400
Qy	1343	ACCGACCTCATGATTTGCTCCGCTGTGCACAGTCCCTGGCAGGTATATGGTTTCT	1402
Db	1401	CCGTGCGTGTGAGCTGCTCCGCCACACAGCAGCCCTGCGCAGGTGTACGGCTTCGT	1460
Qy	1403	TCGGGCTGTCTCTGCAGGTGTGCTGTCTGTGGGTACCGAGCAAAATGAGCG	1462
Db	1461	GGGGGCTTGCCTGCGCGGCTGTGCCCCAGGCCCTCGGGCTCCAGGCACAAAGACG	1520
Qy	1463	CGCCTCTTTAAGAACTTAAAGAGTTTCATCTCGTTGGGAAATACGGCAAGCTATCACT	1522
Db	1521	CGCCTCTCAGGACACCAAGAGTTTCATCTCCTGGGAAGCATGCCAAGCTCTCGT	1580
Qy	1523	GCAGGAATGATGTGGAAGATGAAGTAGAGGATGGCACTGCTCGCAGCAGCCGG	1582
Db	1581	GCAGGAGCTGACGTGGAAGATGAGCGTGGGGAGTGCCTTGGCTGCGCAGAGCCGAG	1640
Qy	1583	GAAGNACGCTGCCCGCTGCAGAGCACCGTCTCAGGAGGAGGATCTCTGCTACGTTCT	1642
Db	1641	GGTGGCTGTGTCCGGCGCAGACACCGTCTGCTGAGGAGATCTTGCCAAAGTTCT	1700
Qy	1643	GTTCTGGCTGATGGACATACAGTGGTACAGCTGCTTAGTGCATCTTTTACATCAGA	1702
Db	1701	GCACCTGGCTGATGAGTGTACGTCTCGAGCTGCTCAGGTCTTTCTTTTATGTCGGA	1760
Qy	1703	GAGCACATTCAGAGAACAGCTCTTCTTACCGTAAGAGTGTGTGGAGCAAGCTGCA	1762
Db	1761	GACCACGTTTCAAAAGACAGGCTCTTTTCTACCGGAAGAGTCTCTGGAGCAAGTTGCA	1820
Qy	1763	GAGCATTTGAGTCAGCAACACTTTGAGAGAGTGGCGTACGGAGCTGTCCAAAGAGA	1822
Db	1821	AAGCAITGGAATCAGACACACTTTGAAGAGGTGCAGCTGGGGAGCTGTCCGAAGCAGA	1880
Qy	1823	GGTCAGGATCACCAGGACACTGGCTAGCCATGCCATCTGCAGACTGCGCTTCATCCC	1882
Db	1881	GGTCAGCAGCATTCGGGAAGCAGGCCGCGCTGCTGAGTCCAGACTCCGCTTCATCCC	1940
Qy	1883	CAAGCCCAAGCGCTCGGCCCATCTGTAACATGAGTTATAGCATGGGTACAGAGCTTT	1942
Db	1941	CAAGCCTGAGGCTCGGCCGATTTGAACTGGACTACGTCGTGGGAGCCAGAACGTT	2000
Qy	1943	GGCGAAGGAAGCAGGCCAGCATTTTACCAGCGTCTCAGACTCTCTTCAGCATGCT	2002
Db	2001	CCGCAGAGAAAGAGGCCAGCGTCTCACTCGAGGTTGAAGGACACTGTTTACGCGTCT	2060
Qy	2003	CAACTATGAGCGGACAAACATCCTCACTTATGGGGCTTCTGTACTGGGTATGAATGA	2062
Db	2061	CAACTAGCGGGCGCGCGCCCGCCCTCTGSGCGCTCTGTGCTGGGCTGGAGCA	2120
Qy	2063	CATCTACAGGACCTGGCGGCTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	2122
Db	2121	TATCCACAGGCGCTGGCGCACTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	2180
Qy	2123	GATGTACTTTTGAAGCAGATGTGACCGGGCTTATGATGCCATCCCCAGGGTAAGCT	2182
Db	2181	GCTGTACTTTTCAAGTGTGATGTGCGGGCGGTACGACACCATCCCCAGGACAGGCT	2240
Qy	2183	GGTGGAGTCTTGCCAAATGATFAGGCACTCGGAGAGCAGTACTGTATCGCCAGTA	2242
Db	2241	CACGGAGTCTATGCCAGCATCA---AAGCCAGAACACGTACTTGGGTGCGTCGTA	2297
Qy	2243	TGCAGTGTTCGGAGAGATACCGAAGCCAAAGTCCACAGTCTTTTAGGAGACAGTCAAC	2302
Db	2298	TGCGGTGTCCAGAAGGCCGCCATGGCAGTCCGGAAGGCTTCAAGAGCCACGCTTC	2357
Qy	2303	CACCTCTCTGACCTCCAGCCATACATGCGGCGAGTTCCTTTAAGCATCTGCAGGATTCAGA	2362
Db	2358	TACCTTGACAGACTCCAGCCGTACATGCGACAGTTCGTGGCTCACCTTCGAG-----GA	2411
Qy	2363	TGCCAGTGACTGAGGAACCTCGTTGTATCGACAGCATCTCTATGATGAGAGCAG	2422

CC tissues or (ii) by measuring the amount of the encoding nucleic acid.
 CC Expression of the nucleic acid encoding the subunit mRNA is confined to
 CC tumour cells, in contrast to the ubiquitous expression of the telomerase
 CC RNA subunit.
 SQ Sequence 2089 BP; 380 A; 641 C; 627 G; 441 T;

Query Match 23.6%; Score 826.8; DB 1; Length 2089;
 Best Local Similarity 74.3%; Pred. No. 1.2e-227;
 Matches 1073; Conservative 0; Mismatches 362; Indels 9; Gaps 2;

1736 CCGTAAAGAGTGTGGAGCAAGCTGCGAGAGTGGAGTGCAGCAACACCTTTCAGAGAGT 1795
 Db 1 CCGGAAGAGTGTGGAGCAAGTGTGCAAGATTTGAAATCAGACAGACTTGAAGAGGT 60
 1796 GCGGCTACGGGAGGTGTCAAGAGAGGAGTTCAGCATCACAGGACACCTTGGCTAGCAT 1855
 Db 61 GCAGCTCGGGAGGTGTGGAAGAGAGGTCAGGAGCATCGGGAAGCAGGCGCGCCT 120
 1856 GCCCATCTGAGAGTGGCTTCATCCCAAGCCCAAGCGCTGCGGCCATTTGTGAACAT 1915
 Db 121 GCTGACGTCCAGACTCCGCTTCATCCCAAGCCTGAGCGGCTGCGGCCGATTTGAACAT 180
 1916 GAGTTATAGCATGGGTACAGAGCTTTGGGCAAGAGGAGGAGCCAGCAGCATTTACCCA 1975
 Db 181 GGACTAGCTGTGGAGCCAGAGCTTCCGAGAGAGAGAGGAGGCGGTCTCACCTC 240
 1976 GCGTCTCAAGACTCTCTTACAGATGCTCACTATGAGGAGGAGCAAAACATCTTCACTTAT 2035
 Db 241 GAGGTGAAGCACTGTTCAGCTGCTCACTAGCAGGCGGCGGCGCGCGCGCTCTCT 300
 2036 GGGTCTTCTCTACTGGTATGATGATCATCTACAGAGCTGCGGCGCTTTGTGCTGCG 2095
 Db 301 GGGCGCTCTGTCTGGGCTTGGAGATATCCAGAGGCTTGGCGCCTTCGTGCTGCG 360
 2096 TGTGCTGCTCTTGGAGCAGACACCCAGGATGATCTTTTAAAGGAGATGTGACCGGGG 2155
 Db 361 TGTGCGGCGCCAGGAGCGCGCTGAGTGTCTTGTCAAGGTGGATGTGAGCGGCGC 420
 2156 CTATGATGCCATCCCGAGGATTCAGATGCTGGTGGAGTGTTCGCAATATGATCAGCACT 2215
 Db 421 GTACGACACCATCCCGAGGAGAGGCTCAGGAGGTCTATCCAGCATCATCA ---AACC 477
 2216 GGAGAGCAGCTACTGTATCCCGCAGTATGCTGTCGAGAGATAGCCAAAGGCCAAGT 2275
 Db 478 CCAGAACAGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 537
 2276 CCACAAGTCTTTAGGAGAGAGGTACCAACCTCTCTGACCTCCAGCAGCATATGAGGCGCA 2335
 Db 538 CCGCAAGGCTTCAAGAGCCAGCTCTTACCTTGACAGACCTCCAGCGGTATATCGGACA 597
 2336 GTTCCTTAAGCATCTGCAGGATTCAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2395
 Db 598 GTTCGTGGCTTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 651
 2396 GCAGAGCATCTCTATGAATGAGAGCAGCAGCAGCTGTTTGTGCTGCTGCTGCTGCTGCT 2455
 Db 652 GCAGAGCTCTCTCTGATGAGGCGCAGGAGTGGCTCTCTGCTGCTGCTGCTGCTGCT 711
 2456 GGTCTACAGTCTGTAAAGATTGCTGACAGGCTGCTATAGCAGTGCAGGCGCATCCCGCA 2515
 Db 712 GTGCCACACGCGGTGCGCATCAGGCGCAAGTCTTACGTCCAGTGCAGGAGGATCCCGCA 771
 2516 GGGCTCCAGCTTATCCAGCCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2575
 Db 772 GGGCTCCATCTTCCAGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 831
 2576 GTTGTGCTGAGTGCAGCGGATGGGTGCTTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2635
 Db 832 GTTGTGCGGAGTTCGCGGAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 891
 2636 GAGCGCTCACTTGGAGCAAGCAAGCAAAACCTTCTCTACAGCAGCTGCTGCTGCTGCTGCT 2695

2892 ACCTTCCAGAGTGTCTTCAAGCTGGGAGACCATCGGAGCAAGCTCCTGCTGCTGCTGCTG 2951
 Db 1252 ACCTTCAACCGCGGCTTCAAGGCTGGGAGCAATGCGTGCAGAACTCTTTGGGCTCTTG 1311
 2952 CGGTTGAAGTGTACGGTCTATTTCTAGACTTGCAGAGTGAACAGCCTCCAGACAGTCTGC 3011
 Db 1312 CGGCTGAAGTGTACAGCTGTTCTGATTTGAGGTGAACAGCCTCCAGAGGCTGTCG 1371
 3012 ATCAATATATCAAGATCTTCTGCTTACGGCTACAGTTCATGATGATGATGATGATGAT 3071
 Db 1372 ACCAATATCAAGATCTTCTGCTGCTGAGGCTACAGTTCATGATGATGATGATGATGAT 1431
 3072 CTTCCCTTGGACGCTGTTAGGAGAACCTTCATCTTTCTGGGATCATCTATCTGAC 3131
 Db 1432 CTTCCATTTATCAGCAAGTTTGGAGAACCCACATTTTCTGCGGTCTATCTGAC 1491
 3132 CAAGCATCTGCTGATCTATCTCTGAGGTCAAGATTCAGGATTCAGATGATGATGATGAT 3189
 Db 1492 ACGGCTCTCTGCTTACTTCTCTGAGGCGGAGGATGCTGCTGCTGCTGCTGCTGCTG 1551
 3189 -----GCCTCTGGCTCTTCTCTGAGGCGGAGGATGCTGCTGCTGCTGCTGCTGCTG 3242
 Db 1552 AAGGCGCGCGCGGCTCTGCTGAGGCGGAGGATGCTGCTGCTGCTGCTGCTGCTGCTG 1611
 3243 CTGCTCAAGCTGCTGCTCTATCTGATCTATCAAAATGCTCTCTGGGAGCTCTGAGGACA 3302
 Db 1612 CTGCTCAAGCTGCTGCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCT 1671
 3303 GCCCAAACTGCTGCTGCGGAGCTCCAGAGGCGGAGGATGATGATGATGATGATGATGAT 3362
 Db 1672 GCCCAGCAGCAGTGTAGTGGAGAGTCCCGGAGCAGCAGTGTGCTGCTGCTGCTGCTG 1731
 3363 GCTGACCCAGCCCTTAAGCAGAGCTTTCAGACCATTTTGGACTAA 3407
 Db 1732 GCCAACCGGCACTGCTTCAAGACTTCAAGACCATCTGAGCTGA 1776

RESULT 9
 ID V72126
 AC V72126 standard; cDNA; 2089 BP.
 DT 24-MAY-1999 (first entry)
 DE Human catalytic telomerase sub-unit cDNA 3' end variant.
 KW Human; catalytic telomerase subunit; therapy; diagnosis; hTC; assay;
 KW modulator; treatment; inhibit; cellular disorder; death; defect; cancer;
 KW ageing; antisense; neoplastic cell; telomerase-related condition;
 KW tumour cell; ss.
 OS Homo sapiens.
 PN W09859040-A2.
 PD 30-DEC-1998.
 PF 09-JUN-1998; E03468.
 PR 14-APR-1998; DE-016496.
 PR 20-JUN-1997; DE-026329.
 PR 26-MAR-1998; DE-013274.
 PA (FARB) BAYER AG.
 PI Hagen G, Siegmund H, Weichel W, Wick M, Zubov D;
 DR WI; 99-081276/07.
 PT New catalytically active subunit of human telomerase - used in the
 PT modulation of telomerase activity, particularly for treating cancer
 PT and ageing
 PS Disclosure; Fig 14: 76pp; German.
 CC This sequence encodes a novel human catalytic telomerase sub-unit
 CC (hTC). The encoded protein can be used in screening assays to identify
 CC modulators of telomerase and to treat or inhibit cellular disorders,
 CC death, defects and/or other pathological processes involving telomerase,
 CC stimulate, inhibit or mimic the activity of the subunit). Antisense
 CC nucleic acids inhibit telomerase action (by binding to specific mRNA),
 CC particularly in neoplastic cells and may be expressed in vivo. Antibodies
 CC and fragments of the protein, used as probes or primers, are used to
 CC diagnose telomerase-related conditions (especially neoplasia) by (i)
 CC detecting abnormal levels of the subunit protein in body fluids or

Mon Aug 16 10:21:24 1999

```
Db 950 TTGTTAGATCGCGGCCACCGGCTATTCCTCCCTGGTGGCGCCCTGCTGGATACCCGG 1009
QY 2832 ACTTTGGAGGTGTTCTGTGACTACTCAGGTTATGCCAGACCTCAATTAAGACGAGCCTC 2891
Db 1010 ACCCTGGAGGTGACAGGAGTACTCTCAGCTATGCCGGACCTCCATCAGAGCCAGTCTC 1069
QY 2892 ACCTTCCAGAGTGTCTTCAAGCTGGGAGACCATGCGGGAACAGCTCTGCTGCTTG 2951
Db 1070 ACCTTCAACCGCGCTTCAAGCTGGGAGACATGCGTCCAACTCTTTGGGCTTG 1129
QY 2952 CGGTGAAGTGCAGGCTATTTCTAGCTTGCAGGTGAACAGCCTCCAGACAGTCTGC 3011
Db 1130 CGGCTGAAGTGCAGGCTATTTCTGATTTGCAGGTGAACAGCCTCCAGACGCTGTC 1189
QY 3012 ATCAATATATAAGATCTTCTGCTTACAGGCTACAGGTTCCATGATGTGATTCAG 3071
Db 1190 ACCAATCTACAAGATCTCTGCTGCTGAGGCTACAGGTTTACGCAATGTGTGCTGAG 1249
QY 3072 CTTCCCTTTGACCGAGTGTAGGAGACCTCAGATTTCTTGGCATCATCTCCAGC 3131
Db 1250 CTCCCATTTTCAGCAAGTTTGAAGAACCCACATTTTCTGCGGCTCATCTCTGAC 1309
QY 3132 CAAGCATCTGCTGCTATGCTATCTTCAAGGTCAAGATCCAGGATGACATTAAG--- 3189
Db 1310 ACGGCTCTCTGCTGCTTACCTCCATCTGAAAGCAAGACGAGGATGTGCTGGGGGCC 1369
QY 3189 -----GCCTCTGGCTCTTCTCTCTGAGCGGCACATTTGCTCTGCTACCAAGCCTTC 3242
Db 1370 AAGGCGCCCGCGCCCTCTGCTCTCGAGGCGGTGCTGCTGCTGCTGCTGCTGCTGCTG 1429
QY 3243 CTGCTCTGAGTGTGCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTG 1489
Db 1430 CTGCTCTGAGTGTGCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTG 1549
QY 3303 GCCCAAAACTGCTGTGCCGGAAGTCCAGAGCGGCATGACCATCTCTTAAAGCTGCA 3362
Db 1490 GCCCAGAGCGAGCTGAGTGGGAGCTCCCGGGAGCGAGCTGCTGCTGCTGCTGCTGCTG 1549
QY 3363 GCTGACCCAGCGCTTAAACAGACGCTTTTCAAGACCATCTCTGCTGCTGCTGCTGCTG 3407
Db 1550 GCCAACCAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1594

RESULT 11
V22380
ID V22380 standard; cDNA; 2176 BP.
AC V22380:
DT 13-AUG-1998 (first entry)
CF Human telomerase reverse transcriptase cDNA clone 712562.
KW Human telomerase reverse transcriptase; hTERT; TRT; diagnosis;
KW prognosis; cell proliferation; cancer; ageing; ribonucleoprotein; ss.
OS Homo sapiens.
EH Key
FI Location/Qualifiers
FT 23..802
FT CDS
FT /tag= a
FT /product= "telomerase reverse transcriptase"
FT
PN GB2317891-A.
PD 08-APR-1998.
PE 01-OCT-1997; 020890.
PR 14-AUG-1997; US-915503.
PR 01-OCT-1996; US-724643.
PR 18-APR-1997; US-844419.
PR 25-APR-1997; US-846017.
PR 06-MAY-1997; US-851843.
PR 09-MAY-1997; US-854050.
PR 14-AUG-1997; US-911312.
PR 14-AUG-1997; US-912951.
PA (GERO-) GERON CORP.
PA (UYTE-) UNIV TECHNOLOGY CORP.
PA Andrews WH, Cech TR, Chapman KB, Harley C, Lingner J,
PI Morin GB, Nakamura T, Harley CB;
DR WPI; 98-171633/16.
DR P-PSDB; W46998.
```

Pure and recombinant human Telomerase Reverse Transcriptase and its variants - are useful in the diagnosis, prognosis and treatment of cell proliferation conditions especially cancer and ageing

Example 1: Fig 18; 387bp; English.
The present sequence represents a human telomerase reverse transcriptase (hTERT) cDNA clone from the present invention. The present invention also describes the following methods: (A) determining whether a test compound is a modulator of hTERT, by detecting the change in hTERT recombinant protein or polynucleotide, on administration of the compound; (B) preparation of recombinant telomerase by contacting a protein with the hTERT RNA or protein in a sample by binding a relevant probe to the sample and detecting the complex formed or in the case of RNA detection, amplifying the product and correlating the presence of complex or amplification product with presence of hTERT in the sample; and (D) increasing the proliferation of a vertebrate cell by increasing hTERT expression; and (E) the use of an agent that causes an increase in cell proliferation to create a medicament that inhibits the proliferation of a cell. A protein preparation of hTERT and the polynucleotide encoding hTERT can be used in the manufacture of medicaments for inhibiting the effect of ageing or cancer. Inhibitors of telomerase activity can be used to treat conditions that are associated with high telomerase activity. A protein preparation of hTERT can also be used in the new methods.

Sequence 2176 BP: 432 A; 679 C; 641 G; 422 T;

Query Match 20.4%; Score 714.4; DB 1: Length 2176;
Best Local Similarity 66.3%; Pred. No. 2,6e-195;
Matches 1185; Conservative 0; Mismatches 401; Indels 200; Gaps 3;

```
QY 1631 GGCTACGTTCTGCTGCTGATGACACATACGTTGGTACGCTGCTTAGTCATCTT 1690
Db 1 GGCAAGTTCTGCTGCTGCTGATGAGTGTACGCTGCTGCTGCTGCTGCTGCTGCT 60
QY 1691 TTACATCAGAGAGACATCTCCAGAGACAGAGCTCTTCTTACCGTAAGAGTGTGT 1750
Db 61 TTATGTCAGGAGAGCCACGTTTCAAAAGAACAGAGCTCTTTTCTACCGAAGAGTGTCTG 120
QY 1751 GAGCAAGCTGCAGAGCATTTGGAGTCAGGCAACACCTTTGAGAGAGTGGCGGTAGGGAGCT 1810
Db 121 GAGCAAGTTGCAAGCATTTGGAATCAGACAGCAGCTTGAAGAGGTGCGAGCTGCGGAGCT 180
QY 1811 GTCAAGAGGAGGTGAGGATCAGGACACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1870
Db 181 GTCGGAAGCAGAGGTTCAGGAGCATCGGGAAGCCAGGCGCCCTGCTGCTGCTGCTGCTGCT 240
QY 1871 GCGCTTCATCCCAAGCCCAAGCGCTCGCGCCCATTTGTGAACATGAGTTATAGCATGGG 1930
Db 241 CCGCTTCATCCCAAGCCCTGAGGCGCTCGCGCCGATTTGGAACATGAGTACGCTGCTGCTGCTG 300
QY 1931 TACCAAGAGTCTTGGCAGAGAGGAGCAGGCGCCAGCATTTACCCAGCGCTCTCAAGACTCT 1990
Db 301 AGCCAGACGTTCCGAGAGAAAGAGGCGGAGCGCTCTCACTTCAGGAGGTGAGGCACT 360
QY 1991 CTTCAAGCTGCTCAACTATGAGCGGCAACAAACATCTTCACTTATGGGTCTTCTGCTACT 2050
Db 361 GTTCAGGCTGCTCAACTACAGCGGCGCGCGCGCGCCCTCTCTGCGGCTCTGCTGCT 420
QY 2051 GGTATGAATGACATCTACAGGACCTTGGCGGCGCTTTGTGCTGCTGCTGCTGCTGCTGCTG 2110
Db 421 GGGCTGGAGATATCCAGAGGCGCTTGGCGCACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
QY 2111 CCAGACACCCAGGATGCTTCTTAAAGCAGATGTGACCGGGGCGCTATGATGCCATGCC 2170
Db 481 CCGCGCGCTGAGCTGTACTTTGTCAAGGTGATGTGACGGGCGCTACGACACCATGCC 540
QY 2171 CCAGGTAAGCTGGTGGAGTCTTCCCAATATGATGAGGACTCGGAGACAGCTGCTGCTGCTG 2230
Db 541 CCAGGACAGGCTCAGGAGGTCTATCGCCAGCATCA---AACCCCAAGACAGCTGCTG 597
QY 2231 TATCCGCGAGTATGCAAGTGTCTCGGAGAGATAGCCAGGCGCAAGTCCACAAGTCTCTTAG 2290
```


CC associated with telomerase and a constituent of active telomerase
CC complex. The products can be used for increasing or reducing the
CC lifespan of cells such as cancer cells or transformed cells. They
CC can also be used in the diagnosis and treatment of malignancies.
CC In addition, cells with a longer lifespan can be transplanted into
CC or grafted onto an individual (e.g. as skin grafts, as systems for
CC delivery of therapeutic proteins, such as hormones and enzymes), to
CC whom they provide therapeutic benefit.
SQ Sequence 3346 BP; 617 A; 1027 C; 965 G; 735 T;

Query Match 18.5%; Score 647; DB 1; Length 3346;
Best Local Similarity 66.7%; Pred. No. 7.7e-176;
Matches 1062; Conservative 0; Mismatches 340; Indels 191; Gaps 2;

QY 1582 GGAGGACCGTGTCCCGCTGACAGACACCGTCTGAGGGAGAGATCCTGGCTACGTTCC 1641
DB 971 GGGTTGGCTGTGTCCGGCCGACAGACCGTCTGCTGAGGAGATCCTGGCCAAAGTTCC 1030
QY 1642 TGTTCGGCTGATGGACACATAGTGTGACAGCTGCTTAGTGTATCTTTTACATCACAG 1701
DB 1031 TGCACTGGCTGATGAGTGTGCTAGCTGTGAGCTGTCTAGGCTCTTTCTTTATGTACGG 1090
QY 1702 AGAGCACATTCACAGAAACAGGCTCTTCTTACCGTAAGAGTGTGTGAGCAAGCTGC 1761
DB 1091 AGACCAGTTTCAAGAAACAGGCTCTTTCTACCGAAGAGTGTCTGGAGCAAGTTGC 1150
QY 1762 AGAGCATTTGAGTACGGCAACACCTTGAGAGAGTGGGCTACGGGAGCTGTCAAGAGG 1821
DB 1151 AAAGCATTTGGAATCAGACAGCATTTGAAGAGGCTGAGCTGCGGAGCTGTGCGAAGCAG 1210
QY 1822 AGGTACAGGATCACAGGACACCTGTGCTAGCCATGCCATCTGCACACTGCGTTTCATCC 1881
DB 1211 AGGTACAGGACATCGGGAAGCCAGGCGCCCTGTGCTGACCTCAGACTCGCGTTTCATCC 1270
QY 1882 CCAAGCCCAAGCCCTCGCGCCCATTTGTGAACATAGTTATAGCATGGTACCAGAGCTT 1941
DB 1271 CCAAGCCTGACGGCTGCGCGCATTTGTGAACATGACTACGCTGTGGAGCCAGAACGT 1330
QY 1942 TGGGAGAGAGGAGGAGGCGCCAGCATTTACCCAGGCTCTCAAGACTCTCTTCAGCATGC 2001
DB 1331 TCGGAGAGAAAGAGGCGCGGCTCTCACTCGAGGCTGAAGGCATCTGTCAGCGTGC 1390
QY 2002 TCAACTATGAGGGGAGAAACATCTCTACCTTATGGGTCTTCTGACTGAGTATGAATG 2061
DB 1391 TCAACTACGAGCGGCGCGCGCCCTCTGGGCGCCCTGTGCTGGGCGCTGCTGGGCGTGGAG 1450
QY 2062 ACATCTACAGGACCTGGCGGCGCTTTGTGCTGCTGTGCTGCTGTGGACACAGACACCA 2121
DB 1451 ATATCCACAGGCGCTGGCGACCTTCGTGCTGCTGGGCGCCAGGACCGCGCGCTG 1510
QY 2122 GGATGTACTTTGTTAAGGCAGATGTACCGGGCGCTATGATGCCATCCCCAGGGTAAGC 2181
DB 1511 AGCTGTACTTTGTTAAGGTGATGTACGGGCGCTATGATGCCATCCCCAGGACAGCG 1570
QY 2182 TGGTGGAGGTTGTGCCAATATGACGAGCCTCGGAGAGCAGTACTCTATCCGCGCAGT 2241
DB 1571 TCACGGAGGTATCGCCAGCATCA ---AACCCAGAACACGCTACTGCTGCTGCTGGT 1627
QY 2242 ATGCAGTGGTCCGGAGAGATACCAAGGCCAAGTCCCAAGTCTTTAGGAGACAGGTCA 2301
DB 1628 ATGCGGTGGTCCAGAGGCGCGCCATGGGACGTCGCGAAGGCCCTTCAAGAGCCAGTC- 1687
QY 2302 CCACCCCTCTCTGACCTCCAGCCATACATGGGGCCAGTCTCTTAAAGCATCTGCAGGATTCAG 2361
DB 1687 ----- 1687
QY 2362 ATGCCAGTGCACAGGAACTCCGTTGTATCCGAGAGCATCTCTATGAATGAGACCA 2421
DB 1687 ----- 1687
QY 2422 GCAGCAGCCTGTTTCACTTCTCTGCACTTCCCTGCGTCACAGTGTGTAAGATTGTTG 2481

DB 1687 ----- 1687
QY 2482 ACAGGTGCTATACGAGTCCAGGCGCATCCCCAGGGCTCCAGGCTATCCACCGTCTCT 2541
DB 1687 -----CTACGTCCAGTCCAGGGATCCCGCAGGGCTCCATCTCTCCAGGCTCTCT 1739
QY 2542 GCAGTCTGTGTTTCGGAGACATGGAGAACAAAGCTGTTTCTGAGGTGCGAGCGGATGGGT 2601
DB 1740 GCAGCCTGTGCTACGCGGACATGGAGAACAAAGCTGTTTTCGGGGGATTCGGCGGAGCGGC 1799
QY 2602 TGTCTTACGTTTGTGTAAGACTTCTTGTGTTGGTACGCGCTCCTTGGACCAAGCAAAA 2661
DB 1800 TGCTCTCTGCTTGTGTAAGATGATTTCTTGTGTTGACACTCACCTCACCCACGCAAAA 1859
QY 2662 CTTTCTCTACGACCGCTGTGTCATGCGGTCTCTGAGTATGGGTGCGATGATAAATCTTGCA 2721
DB 1860 CTTTCTCTACGACCGCTGTGTCATGCGGTCTCTGAGTATGGGTGCGGTGTAATCTTGCGGA 1919
QY 2722 AGACAGTGGTGAACCTTCCCTGTGAGAGCTGTGTACCTGGGTGGTGCAGCTCCATACCAGC 2781
DB 1920 AGACAGTGGTGAACCTTCCCTGTGTAAGAGCAGAGGCGCTGGGTGGCAGCGCTTTGTTCA 1979
QY 2782 TGCTCTCTACGCTGCTGTTTCCCTGGTGGTGGTCTGCTGCTGGACACTCAGACTTTTGGAGG 2841
DB 1980 TGCCGCGCCACGCGCTATTTCCCTGGTGGGCTGCTGCTGGATACCCGACCGCTTGAGG 2039
QY 2842 TGTCTCTGACTACTACTGCTGATGCTGAGAGCTGTGTACCTGGGTGGTGCAGCTCCATACCAGC 2901
DB 2040 TGCAGAGCGACTACTCTGAGCTATGCTGAGAGCTTCCATCAGAGCGCTCTACCTTCAACC 2099
QY 2902 GTGCTTCAAGCTGGGAGACCATGCGGAACAAGCTCTGCTGCGCTCTGCGGTGGAAGT 2961
DB 2100 GCGGCTTCAAGCTGGGAGAACATGCGTGCMAACTCTTTGGGGTCTTGGGGTGAAGT 2159
QY 2962 GTACGGTCTATTTCTAGACTTGCAGGTGAACAGCCTCCAGACAGTCTGCAATCAATATAT 3021
DB 2150 GTACAGCGCTGTTTCTGAGTTTGCAGGTGAACAGCCTCCAGACGCTGTGCAACACATCT 2219
QY 3022 ACAAGATCTTCTGCTTCCAGCGCTACAGGTTCCATGCTGCTGCTGCTGCTGCTGCTGCTG 3081
DB 2220 ACAAGATCTTCTGCTGCGAGCGTACAGGTTTACAGGATGTTGCTGCGAGCTCCCATTTT 2279
QY 3082 ACCAGCTGTTAGGAAGAACCTCACTTCTTCTGGGCATCATCTCCAGCGCAAGCATCCT 3141
DB 2280 ATCAGCAAGTTTGAAGAACCCACATTTTCTGCGCTCATCTCTGACACGCGCTCCC 2339
QY 3142 GCTGCTATGCTATCTGAGGTTCAAGATCCAG 3174
DB 2340 TCTGCTACTCTCTGAAAGCCAGAACGCGAG 2372

RESULT 14
V72125

ID V72125 standard; cDNA; 3972 BP.

AC V72125, 1999 (first entry)

DE Human catalytic telomerase sub-unit RACE-generated cDNA #4.

KW Human; catalytic telomerase subunit; therapy; diagnosis; hTC; assay;

KW modulator; treatment; inhibit; cellular disorder; death; defect; cancer;

KW ageing; antisense; neoplastic cell; telomerase-related condition; RACE;

KW tumour cell; ss.

OS Homo sapiens.

PN W09859040-A2.

PD 30-DEC-1998.

PF 09-JUN-1998; E03468.

PR 14-APR-1998; DE-016496.

PR 20-JUN-1997; DE-026329.

PR 26-MAR-1998; DE-013274.

PA (FARB) BAYER AG.

PI Hagen G, Siegmund H, Weichel W, Wick M, Zubov D;

DR WPI; 99-081276/07.

PT New catalytically active subunit of human telomerase - used in the modulation of telomerase activity, particularly for treating cancer

PT and ageing

PS Example 6; Fig 11; 76pp; German.

CC This sequence represents a RACE-generated fragment of a novel human
CC catalytic telomerase sub-unit (hTC). The hTC protein can be used
CC in screening assays to identify modulators of telomerase and to treat or
CC inhibit cellular disorders, death, defects and/or other pathological
CC processes involving telomerase, particularly cancer and ageing (also
CC suitable for this are agents that stimulate, inhibit or mimic the
CC activity of the subunit). Antisense nucleic acids inhibit telomerase
CC action (by binding to specific mRNA), particularly in neoplastic cells
CC and may be expressed in vivo. Antibodies and fragments of the protein,
CC used as probes or primers, are used to diagnose telomerase-related
CC conditions (especially neoplasia) by (i) detecting abnormal levels of
CC the subunit protein in body fluids or tissues or (ii) by measuring the
CC amount of the encoding nucleic acid. Expression of the nucleic acid
CC encoding the subunit mRNA is confined to tumour cells, in contrast to
CC the ubiquitous expression of the telomerase RNA subunit.
SQ Sequence 3972 BP; 701 A; 1287 C; 1264 G; 720 T;

Query Match 16.7%; Score 582.8; DB 1; Length 3972;

Best Local Similarity 63.1%; Pred. No. 2.4e-157;

Matches 1046; Conservative 0; Mismatches 522; Indels 90; Gaps 6;

QY 23 CCCGGCCTTGAGCAACAATGACCCCGCTCTCTGTTGGCCCGCGTGGCTCTCTGCTGGC 82
DB 67 CCCGGCCACCCCGCATGCGCGCTCCCGCTGCGAGCCGCTGCTCTCTGCTGGC 126
QY 83 CAGCGATACCGGAGGTGGCGCTGGCAACCTTTTGGCGCGCTGGGGCCCGAGGG 142
DB 127 CAGCCACTACCGGAGGTGGCGCTGGCCACCTTCTGGCGCGCTGGGGCCCGAGG 186
QY 143 CAGCGGCTTGTCAACCCCGGGACCCGCAAGATCTACCGCACTTTGGTTGGCCAAATGCCT 202
DB 187 CTGGCGCTGTGACGCGCGGGACCCGCGCGCTTCCGCGCTGTGGCGCCAGTGCCT 246
QY 203 AGTGTGATGACATGGGGCTCACAGCTTCACCTGGCGACCTTTCTTCCACAGGTGTC 262
DB 247 GGTGTGCTGCTGGAGCAGCAGCGCCCGCCCGCGCGCGCTTCCGCCAGGTGTC 306
QY 263 ATCCCTGAAGAGCTGTGGCGAGGTGTGACAGACTCTGCGAGCGCAACGAGAGAA 322
DB 307 CTGCTGAAGAGCTGTGGCGAGTCTGACAGGCTGTGCGAGCGCGCGAGAA 366
QY 323 CGTGTGCTTTTGGCTTTGAGCTGCTTTAACGAGCCAGAGCGGGCTCCCATGGGCTT 382
DB 367 CGTGTGCTTGGCTTTCGCTGCTGAGCGGGCCCGCGGGCCCGCCCGAGGCTT 426
QY 383 CACTAGTAGCTGCTAGCTACTTCCCAACACTGTTATTAGACCTTGGCTGCTAGTGG 442
DB 427 CACCAACAGCTGCGCAGCTACCTGCCCAACACAGGTGACCGACACTGCGGGGGAGCG 486
QY 443 TGCATGGATGTACTGTTGAGCCGAGTGGGGACGACCTGTGTGTTACCTGCTGGCACA 502
DB 487 GCGTGGGGCTGTGCTGCGCGCTGGCGACGACGCTGTGTTACCTGCTGGCAGC 546
QY 503 CTGTGCTTTTATCTCTGGTGGCCCGCCAGCTGTGCTTACAGGTGTGTGGTCTCCCT 562
DB 547 CTGCGCGCTTCTTGTGCTGGTGGCTCCAGCTGCGCTTACCAGTGTGTGCGGGCCGCT 606
QY 563 GTACCAAAATTTGCGCACACGATATCTGGCCCTTGTGTCGCTAGTTACAGGCCAC 622
DB 607 GTACAGCTCGGCGCTGCGCACTAGCGCGCGCCCGCCGCGCAGCTAGTGGACCCGAG 666
QY 623 CGGACCGCTGGGAGGAATTTCACTAACCTTAGTGTCTTACACAGATCAAGAGCAGTAG 682
DB 667 GCGTCTGGATGC-----GAAAGCGCTGGAACCATAG 699
QY 683 TCGCAGGAAGACCGCAACCCCTGGCTTGCATCTCGAGGTACAGAGGCACTGAG 742
DB 700 CGTAGGAGGCGGGGTCTCCCTTGGCTGCGAGCCCGGTGGAGAGGCGCGGGG 759
QY 743 TCTCACCAGTACAAAGTGTGCTTACAGTTAAGAGGCCAGATGCTATCTGCTCCCGAGAGT 802

DB 760 CAGTCCAGCCGAAAGTCTGCGTTGCCAAGAGGCCAGGCTGCGCTGCCCCGTAGCC 819
QY 803 GGAGGAGGACCCACAGCAGAGTCTTACCAACCCCATCATGAGCAATCATGGTGCACAG 862
DB 820 GGAGCGGAGCCGCTTGGCAGGGTCTCTGGGCCCCACCCGCGAGCGGTGGACCGAG 879
QY 863 TCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 913
DB 880 TGACCGTGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 939
QY 914 AGAAAGAGTCTGACCTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 970
DB 940 GGGTGGCTGCTTGGCAGCGGCTTCCACCCATCCGCTGGGCCCGCCAGCAGCAGCGGG 999
QY 971 CTCACATCTC---TGTCTGCTACACCCCGCCAAATGCTTTTACAGCTCAGGCCATTTAT 1027
DB 1000 CCCCCATCCATCGCGGCCACCGCTCCCTGGGACAGCGCTTGTCCCGCGGTGTACGC 1059
QY 1028 TGAGACCAAGATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1087
DB 1060 CGAGACCAAGCACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1116
QY 1088 CCTACTCAGCAACCTCCAGCCTAACTTGACTGGGGCCAGAGACTGGTGGAGATCATCTT 1147
DB 1117 CTTACTCAGCTCTCTGAGCCCGAGCTGACTGGCGTCTGGAGCTCTGAGACCATCTT 1176
QY 1148 TCTGGCTCAAGSCCTTAGACATCAGGACACTCTCTGAGGACACACCCGCTTATCGGCTG 1207
DB 1177 TCTGGTTCAGSCCTTGGATGCCAGGACTCCCGCAGCTTCCCGCGCTGCCCCAGCG 1236
QY 1208 ATACTGGAGATCGGCGCTGTTCCACAGCTGCTGGTGAACCATGACAGTGCACAATA 1267
DB 1237 CTACTGGCAATCGGCGCTGTTCTTGGAGCTGCTGGGAACCCAGCCAGTGGCCCCCTA 1296
QY 1268 TGTCAAGCTCTCAGTTCACATTCAGAGTTTCGAACAGCAAAACCAAGGTGACAGATGC 1327
DB 1297 CGGGGTGCTCTCAAGACGACTGCCCGCTGCGAGCTGCGGTCAACCCAGCAGCGGTGT 1356
QY 1328 CT-----TGACACAGCC 1342
DB 1357 CTGTGCGGGAGAACCCCGAGGCTCTGTGGCGGCCCGCGAGGAGGAGACACAGCC 1416
QY 1343 ACCGCACTCATGATTTGCTCGCTGCACAGCTGCTGCGAGTATATGTTTCT 1402
DB 1417 CCGTGGCTTGGTGGCTGCTCGCAGACACAGCCCTGCGAGGTGTACGGCTCGT 1476
QY 1403 TCGGGCTGCTCTGCAAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1462
DB 1477 GCGGCTGCTGCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1536
QY 1463 CCGTCTTTTAAAGTAAAGTTCATCTGTTGGGGAATACGGCAAGCTATCACT 1522
DB 1537 CCGTCTTCTCAGGAACACCAAGAGTTTCATCTCCCTGGGGAAGCATGCCAAGCTCTGCT 1596
QY 1523 GCAGAACTGATGGAAGATGAAAGTAGAGATTGCACTGGCTGCGCAGCAGCCCGG 1582
DB 1597 GCAGAGCTGACGTGGAAGATGACGCTGCGGACTGCGCTTGGCTGCGCAGGAGCCAGG 1656
QY 1583 GAAGAGCGTGTCCCGCTGCGAGAGCAGCTGCTGAGGG 1620
DB 1657 TGAGGAGGTGTGGCTGCGAGGGGCCCGAGGCCCGAGAG 1694

RESULT 15

V16979

ID V16979 standard; DNA; 4335 BP.

AC V16979;

DT 13-AUG-1998 (first entry)

DE Human telomerase reverse transcriptase genomic clone.

KW Human; telomerase reverse transcriptase; hTERT; TERT; diagnosis;

OS prognosis; cell proliferation; cancer; ageing; ribonucleoprotein; ss.

OS Homo sapiens.

Mon Aug 16 10:21:24 1999

QY 1359 TTGCTCCGCTGACAGAGTCCCTGGCAGGTATATGGTTTTCTTGGGCTGTCTGTC 1418
DB 3950 CTGCTCCGCCAGACAGAGCCCTTGGCAGGTGTACGGCTTCGTGCGGGCTGCTGCGC 4009
QY 1419 AAGGTGGTGTCTGCTAGTCTCTGGGGTACCAGGCACAAATGAGCGCGCTTCTTTAAGAAC 1478
DB 4010 CGGCTGGTGCCCCCAGGCTCTGGGGCTCCAGGCACAAACGACGCCCTTCTCAGGAAC 4069
QY 1479 TTAAGAAAGTTTCATCTCGTTGGGGAATACGGCAAGCTATCACTGCAGGAACGTATGTGG 1538
DB 4070 ACCAAGAAGTTTCATCTCCCTGGGGAAGCATGCCAAGCTCTGCTGCAGGAGCTGACGTGG 4129
QY 1539 AAGATGAAGTAGAGGATGCCACTGGCTCCGACAGCGCCGGGAAGACCGTGTCCCC 1598
DB 4130 AAGATGAGCGTCGGGAGTGGCTTGGCTGCGCAGGAGCCAGGTGAGGAGGTGGTGCC 4189
QY 1599 GCTGCAGAGCACCGTCTGAGGG 1620
DB 4190 GTCAGGGCCCCAGGCCCCAGAG 4211

Search completed: August 13, 1999, 13:03:09
Job time: 6142 sec

Human Immunodeficiency Virus
Tetrahymena macronucleus

ALIGNMENTS

August 13, 1999, 21:18:14 ; Search time 17.01 Seconds
(without alignments)

AC	W47008;
DT	13-AUG-1998 (first entry)
DE	Glutathione-S-transferase and hrrT fusion protein 8.
KW	Human; telomerase reverse transcriptase; hTERT; hTAT; diagnosis;
OS	prognosis; cell proliferation; cancer; ageing; ribonucleoprotein.
KS	Synthetic.

BLOSUM62

OS Homo sapiens.

FT	Misc_difference	22.	.23	/note= "enterokinase cleavage site"
FT				
FT				

PN	GB2317891-A:	
PD	08-APR-1998:	
PE	01-OCT-1997:	020890.
PF	14-AUG-1997:	US-915503.
PG	01-OCT-1996:	US-724643.
PH	18-APR-1997:	US-844419.
PI	25-APR-1997:	US-846017.
PJ	06-MAY-1997:	US-851843.
PK	09-MAY-1997:	US-854050.
PL	14-AUG-1997:	US-911312.
PM	14-AUG-1997:	US-912951.
PN	(GERO-) GERON CORP.	
PO	(UYTE-) UNIV TECHNOLOGY CORP.	
PP	Andrews GB, Cach TR, Chapman KB, Harley C, Lingner J,	
PQ	Morin GH, Nakamura T, Harley CB;	
PR	WPI: 98-171633/16.	
PS	Pure and recombinant human Telomerase Reverse Transcriptase and its	
PT	variants - are useful in the diagnosis, prognosis and treatment of	
PU	cell proliferation conditions especially cancer and ageing	
PV	Example 6; Page 234-235; 387pp: English.	
PW		

-724-

25-APR-1997; US-846017.
PR 06-MAY-1997; US-851843.
PR 09-MAY-1997; US-854050.
PR 14-AUG-1997; US-911312.
PR 14-AUG-1997; US-912951.
PA (GERO-) GENON CORP.
PA (UYTE-) UNIV TECHNOLOGY CORP.
PI Andrews WH, Cech TR, Chapman KB, Harley C, Lingner J,
PI Morin GB, Nakamura T, Harley CB;
PI WPI: 98-171633/16.
DR Pure and recombinant human Telomerase Reverse Transcriptase and its
PT variants - are useful in the diagnosis, prognosis and treatment of
PT cell proliferation conditions especially cancer and ageing
PT Example 6; Page 234-235; 38pp; English.
PS The present sequence represents a fusion protein from an example
CC of the present invention which describes human telomerase reverse
CC transcriptase (hTERT). The present invention also describes the
CC following methods: (A) determining whether a test compound is
CC a modulator of hTERT, by detecting the change in hTERT recombinant
CC protein or polynucleotide, on administration of the compound;
CC (B) preparation of recombinant telomerase by contacting a protein
CC preparation of hTERT with a telomerase RNA component; (C) detection
CC of the hTERT RNA or protein in a sample by binding a relevant
CC probe to the sample and detecting the complex formed or in the case of
CC RNA detection, amplifying the product and correlating the presence of
CC complex or amplification product with presence of hTERT in the sample;
CC and (D) increasing the proliferation of a vertebrate cell by increasing
CC hTERT expression; and (E) the use of an agent that causes an increase in
CC cell vertebrate cell proliferation to create a medicament that inhibits
CC ageing. A protein preparation of hTERT and the polynucleotide encoding
CC hTERT can be used in the manufacture of medicaments for inhibiting the
CC effect of ageing or cancer. Inhibitors of telomerase activity can be
CC used to treat conditions that are associated with high telomerase
CC activity. A protein preparation of hTERT can also be used in the new
CC methods.
SQ Sequence 1189 AA.

Query Match 59.4%; Score 3505; DB 1; Length 1189;
Best Local Similarity 62.4%; Pred. No. 0;
Matches 719; Conservative 122; Mismatches 260; Indels 52; Gaps 13;

QY	1	MTRAPRC	PAVRSLL	RSRYRE	VWPLAT	FVRLRG	PEGRRL	VQPGDP	KYIYTL	VAQCLV	CMHW	60
Db	58	MPRAPRC	AVRSLL	RSRYRE	VYLPLAT	FVRLRG	PQGWRL	VQRGDP	AAFRAL	VAQCLV	CPVW	117
QY	61	GSQPPDL	SFHOVSS	KLKELV	ARYVOR	LACERN	ERNVLA	FGFELL	NFARGG	PPMAFT	SSVR	120

```

Db 118 DARPPAAPSFQVSCLEKELVAVLQRLCERGAKNVLAFCGALLDGAAGPPPEAFTTSVR 177
      ::||| || || |||||:||||| :||||| ||: ||||| ||:|
121 SYLNTVETLRVSGAWMLLSRVGDDLLVLLAHALYLLVPPSCAYQVCGSPYQICA 180
      ||||| : || ||| ||| |||||:||||| |||||:| ||||| |||||:|
178 SYLNTVTDALRSGAWMLLRVGGDDVLVHLLARCALFVLVAPSCAYQVCGPPLQLGA 237
      ||||| : || ||| ||| |||||:||||| |||||:| ||||| |||||:|
181 TTDWPSVASYRTPRVGRNLTFRLOQIKSSRSQAEKPLALPSRGTKRHLSTLSTS 240
      ||||| : || ||| ||| |||||:||||| |||||:| ||||| |||||:|
238 ATOARPPHAS-GPRRLG-----CERAWNHSVREAGVPLGAPGARRRGGASRS 288
      :||| : || ||| ||| |||||:||||| |||||:| ||||| |||||:|
241 VPSAKKARCVPRVRESEGP-----HRQVLPTPSKSW-VPSPARSEVPTAEKLSK 292
      :||| : || ||| ||| |||||:||||| |||||:| ||||| |||||:|
289 LPLKPRRGAAPERTPTVGQSWAHPGRTGRGDFCVSPAR-----PAEETSLE 343
      :||| : || ||| ||| |||||:||||| |||||:| ||||| |||||:|
293 GKYSLSLS-GSVCKKHKPSSTLSLPPRONAFQLRP-FIETRHFLYSRGDOERLNPSF 350
      :||| : || ||| ||| |||||:||||| |||||:| ||||| |||||:|
344 GALSGRHSPVGRQHAGPPSTSRPRWDTPCPVYAEKTHFLYSSGD-KEQLRPSF 402
      :||| : || ||| ||| |||||:||||| |||||:| ||||| |||||:|
351 LLSNLQPLTGAARLVEIFLGSRPRTSGPLCRTHLSRBYQMRPLFOQLLVNHAECQY 410
      :||| : || ||| ||| |||||:||||| |||||:| ||||| |||||:|
403 LLSLSRPLTGAARLVETIFLGSRPMPGTPRRLPRLPQRYQMRPLFLELLGNHAQCPY 462
      :||| : || ||| ||| |||||:||||| |||||:| ||||| |||||:|
411 VRLSRHCRFTANQVTDAL-----NTSPHMLDLRLHSSPKQVY 452
      :||| : || ||| ||| |||||:||||| |||||:| ||||| |||||:|
463 GVLLKTHCPLEAA---VTPAAGVAREKPGQSVAAPEEDTDRRLVOLLRQHSSPMQVY 519
      :||| : || ||| ||| |||||:||||| |||||:| ||||| |||||:|
453 GFLRACLKVVYASLWCTRNERFFENLAKFISLGRYKLSLQELMKMKVEDCHWLS 512
      :||| : || ||| ||| |||||:||||| |||||:| ||||| |||||:|
520 GFVPACLRLVPPGLGWSRNERFLENTKFFISLGHAKLSLQELTWKMSVRDCAWLRR 579
      :||| : || ||| ||| |||||:||||| |||||:| ||||| |||||:|
513 SPGRDVPAAEHLRERILATFLWMDTVYVQLLRFFYTESTFOKNLFFYRKSWS 572
      :||| : || ||| ||| |||||:||||| |||||:| ||||| |||||:|
580 SPGVGVPAEHLREILAKFLHLSVYVVELLRFYVETTFQKNLFFYRKSWS 639
      :||| : || ||| ||| |||||:||||| |||||:| ||||| |||||:|
573 KLSIGVROHLRYRLSLOEVRHODTWLAMPICRLFIKPNGLRPTVNNYSMGT 632
      :||| : || ||| ||| |||||:||||| |||||:| ||||| |||||:|
640 KLSIGIRQHLRVLRLSLOEVRHODTWLAMPICRLFIKPNGLRPTVNNYSMGT 699
      :||| : || ||| ||| |||||:||||| |||||:| ||||| |||||:|
633 RALGRKQAOFTORLKTFLSMLNERTKPHLMGSSVGLMNDIYRTWRAFLVRALDQ 692
      :||| : || ||| ||| |||||:||||| |||||:| ||||| |||||:|
700 RTFREKRAELTSRVKALFVNLVYERARRPGLLGASVGLGDDIHRARWTFVLVRAQDP 759
      :||| : || ||| ||| |||||:||||| |||||:| ||||| |||||:|
693 TPRMVFYKADVTGAYDAIPQGLVEVYANMRHSEYICRQYAVVRSDGQVHKFSRR 752
      :||| : || ||| ||| |||||:||||| |||||:| ||||| |||||:|
760 PPELYFVKVDVTGAYDIPQDLRTEVIAIITK-PONTYCVARYAVQKAAHGHVRKAFKS 818
      :||| : || ||| ||| |||||:||||| |||||:| ||||| |||||:|
753 QVTLTSLDQYMGOFKHLQSDASALNSVVIQSSISMNESSSLDFLFLHFLRHSVVK 812
      :||| : || ||| ||| |||||:||||| |||||:| ||||| |||||:|
819 HVSTLTDQYMGOFVAVHLOET--SPLDADVIEQSSLEASSGLDFVFLRFMCHAVR 876
      :||| : || ||| ||| |||||:||||| |||||:| ||||| |||||:|
813 IGDCRYTCQCGIPQSSLSLTLCSLCFGDMENKLFVAVQVRDGLLRFVDDPLVTPHLDQ 872
      :||| : || ||| ||| |||||:||||| |||||:| ||||| |||||:|
877 IRGKSYVQCGIPQSSLSLTLCSLCYGDMEKLFAGIRDRDGLLRLVDDPLVTPHLDQ 936
      :||| : || ||| ||| |||||:||||| |||||:| ||||| |||||:|
873 AKTELSTLVHGVPEGCMINLQTVNFPVPEGLTGGAPYOLPAHCLFPWCGLLDTOT 932
      :||| : || ||| ||| |||||:||||| |||||:| ||||| |||||:|
937 AKTELTLVRGVPEYGVGNLRLKTVNFPVPEDEALGGTAFVOMPAHGLFPWCGLLDTOT 996
      :||| : || ||| ||| |||||:||||| |||||:| ||||| |||||:|
933 LEVFCDSYGAQTSIKTSLTQSVFQKAGTMRNKLVSRLKCHGLFLDLQVNSLQTVCI 992
      :||| : || ||| ||| |||||:||||| |||||:| ||||| |||||:|
997 LEVQSDYSYARTSIRASLTNRGFKAGNRMRKLFGLVRLKCHSLFLDLQVNSLQTVCI 1056
      :||| : || ||| ||| |||||:||||| |||||:| ||||| |||||:|
993 NIYKFLQAVRFHACVLTQLPDQVRNKLTFGLIISQASCCYAILKVNPGMTLKAAS 1052
      :||| : || ||| ||| |||||:||||| |||||:| ||||| |||||:|
1057 NIYKILLQAVRFHACVLTQLPDQVRNKLTFGLIISQASCCYAILKVNPGMTLKAAS 1116
      :||| : || ||| ||| |||||:||||| |||||:| ||||| |||||:|
1053 GS---FPEAAHLCYQAFLLKLAHSYVYKCLLGLPLTAOKLCKRLPEATMTILKAA 1109
      :||| : || ||| ||| |||||:||||| |||||:| ||||| |||||:|
1117 GAAGPLPSEAVQWMLCHQAFLLKLAHSYVYKCLLGLPLTAOKLCKRLPEATMTILKAA 1176
      :||| : || ||| ||| |||||:||||| |||||:| ||||| |||||:|
1110 DPALSTDFQTLTD 1122
      :||| : || ||| ||| |||||:||||| |||||:| ||||| |||||:|

```

```

Db 1177 NPALPSDFKTILD 1189

RESULT 2
W46957
ID W46957 standard; Protein; 1132 AA.
AC W46957;
DE Human telomerase reverse transcriptase.
KW Human; telomerase reverse transcriptase; hTERT; TERT; diagnosis;
OS Homo sapiens.
PN GB2317891.A.
PD 08-APR-1998. 020890.
PE 01-OCT-1997; US-915303.
PR 14-AUG-1997; US-724643.
PR 01-OCT-1996; US-724643.
PR 18-APR-1997; US-844419.
PR 25-APR-1997; US-846017.
PR 06-MAY-1997; US-851843.
PR 09-MAY-1997; US-854050.
PR 14-AUG-1997; US-911312.
PR 14-AUG-1997; US-912951.
PA (GERO-) UNIV TECHNOLOGY CORP.
PI Andrews WH, Czech TR, Chapman KB, Harley C, Lingner J,
PI Morin GB, Nakamura T, Harley CB;
PI N-PSDB; V22379.
PT Pure and recombinant human Telomerase Reverse Transcriptase and its
PT variants - are useful in the diagnosis, prognosis and treatment of
PT cell proliferation conditions especially cancer and ageing
PT claim 3; Fig 17; 387pp; English.
PS The present sequence represents human telomerase reverse transcriptase
PS (hTERT), which is a ribonucleoprotein. The present invention also
PS describes the following methods: (A) determining whether a test compound
PS is a modulator of hTERT, by detecting the change in hTERT recombinant
PS protein or polynucleotide, on administration of the compound; (B)
PS preparation of recombinant telomerase by contacting a protein
PS preparation of hTERT with a telomerase RNA component; (C) detection of
PS the hTERT RNA or protein in a sample by binding a relevant probe to the
PS sample and detecting the complex formed or in the case of RNA detection,
PS amplifying the product and correlating the presence of complex or
PS amplification product with presence of hTERT in the sample; and (D)
PS increasing the proliferation of a vertebrate cell by increasing hTERT
PS expression; and (E) the use of an agent that causes an increase in cell
PS vertebrate cell proliferation to create a medicament that inhibits
PS ageing. A protein preparation of hTERT and the polynucleotide encoding
PS hTERT can be used in the manufacture of medicaments for inhibiting the
PS effect of ageing or cancer. Inhibitors of telomerase activity can be
PS used to treat conditions that are associated with high telomerase
PS activity. A protein preparation of hTERT can also be used in the new
PS methods.
SQ Sequence 1132 AA;

Query Match 59.4%; Score 3505; DB 1; Length 1132;
Best Local Similarity 62.4%; Pred. No. 0;
Matches 719; Conservative 122; Mismatches 260; Indels 52; Gaps 13;

```

```

QY 1 MTRAPRCPAYRSLRSRYREVWPLATFVRRLGPEGRRLVQPGDKYRTLVACLCVCMHW 60
      :||| : || ||| ||| |||||:||||| |||||:| ||||| |||||:|
      1 MPRAPRCRAVRSLLRSYREVWPLATFVRRLGPEGRRLVQPGDKYRTLVACLCVCMHW 60
      :||| : || ||| ||| |||||:||||| |||||:| ||||| |||||:|
QY 61 GSOPPPADLSFHQVSSLKELVARVQRLCERNERNVLAFCGALLDGAAGPPPEAFTTSVR 120
      :||| : || ||| ||| |||||:||||| |||||:| ||||| |||||:|
      61 DARPPAAPSFQVSCLEKELVAVLQRLCERGAKNVLAFCGALLDGAAGPPPEAFTTSVR 120
      :||| : || ||| ||| |||||:||||| |||||:| ||||| |||||:|
QY 121 SYLNTVETLRVSGAWMLLSRVGDDLLVLLAHALYLLVPPSCAYQVCGSPYQICA 180
      :||| : || ||| ||| |||||:||||| |||||:| ||||| |||||:|
      121 SYLNTVTDALRSGAWMLLRVGGDDVLVHLLARCALFVLVAPSCAYQVCGPPLQLGA 180
      :||| : || ||| ||| |||||:||||| |||||:| ||||| |||||:|
QY 181 TTDWPSVASYRTPRVGRNLTFRLOQIKSSRSQAEKPLALPSRGTKRHLSTLSTS 240
      :||| : || ||| ||| |||||:||||| |||||:| ||||| |||||:|

```



```

Db 181 ATQARPPPHAS -GPRRLG-----CERAWNSVRAGVPLGLPAPGARRRGGSSASRS 231
QY 241 VPSAKKARCYVPVRVEGP-----HRQVLPSPGKSW-VPSPARSPVPTAEKDLSSK 292
Db 232 LPLKPRRARGAAPEPTPVGCGSWAHPRGRGSDRGFCVVSAPAR-----PAEATSLE 286
QY 293 GKVSLSLS-GSVCKKHPSTSLSPRONAFQLRP-FIETRHFLYSRGDGQERLNPSF 350
Db 287 GALSOTRSHSPSVGRQHAGPPTSRSRPRPDWTPCPVYAEATHFLYSSGD-KEQLRPSF 345
QY 351 LLSNLOPLNLTGARRLVEIFLGSRRPTSGPLCTRHLSSRYWQRPFLQOLLVNHAEQY 410
Db 346 LLSLRPSLTGARRLVETIFLGSRRPMPGTPRRLPLRQYQWQRPFLLELGNHACQY 405
QY 411 VRLLSHCRFTANQOQVTDAL-----NTSPHMLDLRLHSSPWQY 452
Db 406 GVLLKTHCPRAA---VTPAAGVAREKPGSVAAPEEDTPRRLVQLLRHSSPWQY 462
QY 453 GFLRACLVKVSASLWCTRHNERRERFKNKKFISLKGKGLSLQELMKWVKYEDCHLRS 512
Db 463 GFVRACLRRLVPPGLMGSRHNERLNTKFFISLKGKHALSLQELTWKMSVRDCAWLR 522
QY 513 SPGRDVPAAEHLRERILATFLWMDTVVOLLRSFFYITESTFQKNLFYFRKSVMS 572
Db 523 SPGVGCVPAAEHLRERILAKFLHLMVSVVVELLSFFYVETTFQKNLFYFRKSVMS 582
QY 573 KLOSIGVQHLRVLRELISOEVRHODTWLAMPICRLRFPKPNGLRPVNMYSMT 632
Db 583 KLOSIGIRHLKRVOLRELSAEVRHREARPAULTSRURFPKPDGLRPVNMVYVGA 642
QY 633 RALGRKQAHQFQRTUKTLFSLMNVERTKPHLMGSSVLGMDIYRTWRAFVLRLALDQ 692
Db 643 RTRREKRAELTSRYKALFSLVNLVNERARRPGLGASVLGDDIHRNRTFVLVRAQDP 702
QY 693 TPRLMYKADVTGAYDAIPOGKLVEVANNIRHSESTCIROYAVVRDSDQVHKSPRR 752
Db 703 PPELYPKVDVTGAYDTIPQDRLETVIASIIR-PONTYCVRRYAVYVAAHGHVRKAFKS 761
QY 753 QVTTLSLQPYMGQFLKHLQDSASALRNSVVIEQISIMNESSLSLDFELHFLRHSVVK 812
Db 762 HVSTLTLDQPYMRQFVAHLQET--SPLRDVAVVIEQSSLSNEASSGLDFVFRFCHHAVR 819
QY 813 IGDRCYTOCOGIPGSSLSLTLCSLCFGDMENKLPFAEVQRDGLLRLFRVDDFLVTPHLDQ 872
Db 820 IRGKSYVOCOGIPGSSLSLTLCSLCYGDMEKLFAGIRRDGLLRLVDDFLVTPHLDQ 879
QY 873 AKTFSTLVHGVPEYGCMLNLOKTVNPEPVGTLGGAPYQLPAHCLFPWCGLLDTQT 932
Db 880 AKTFSTLVHGVPEYGCMLNLOKTVNPEPVGTLGGAPYQLPAHCLFPWCGLLDTQT 939
QY 933 LEVFCDSYGAQTSKTSLSLTSQSVFKAGKTMNRKLLSVLRKCHGLFLDLQVNSLQTVCI 992
Db 940 LEVQSDYSYARTSIRASLTNRGFKAGNRMRRLFGVLRKCHSLFLLDLQVNSLQTVCI 999
QY 993 NIYKIFLQAYRFHACVQLPDDQVRKNTLFFGLGISLSSOACCYAILKVNPGMTIKAS 1052
Db 1000 NIYKIFLQAYRFHACVQLPDDQVRKNTLFFGLGISLSSOACCYAILKVNPGMTIKAS 1059
QY 1053 GS---FPPEAHNLWCYQAFLLKLAHSYIKCLIGLPLRTAQKLLCKRLKPEATWILKAAA 1109
Db 1060 GAAGPLPSEAVQWLCHQAFLLKLRHRTVYVPLGLSLRTAQKLLCKRLKPEATWILKAAA 1119
QY 1110 DPALSTDFTQILD 1122
Db 1120 NPALPSDFKILD 1132

```

RESULT 3

W61350

ID W61350 standard; Protein: 1154 AA.

AC W61350;

DT 12-OCT-1998 (first entry)

```

DE Human telomerase protein 2 (TP2).
KW TP2; human; telomerase protein 2; cancer; AIDS; ageing; therapy;
OS Homo sapiens.
PN WO9821343-A1.
PD 22-MAY-1998.
PF 13-NOV-1997; U21248.
PR 16-OCT-1997; US-951733.
PR 15-NOV-1996; US-871189.
PR 11-JUN-1997; US-873039.
PA (AMGE-) AMGEN CANADA INC.
PI Harrington LA, Robinson MO;
DR WI; 98-297946/26.
DR N-PSDB; V27876.
PT New nucleic acid encoding human telomerase protein-2 - used for
PT regulating telomerase activity, e.g. for treating cancer or acquired
PT immune deficiency syndrome
PS Claim 1e; Fig 9; 150pp; English.
CC This polypeptide comprises human telomerase protein 2 (TP2), a
CC novel protein of the telomerase complex. Its amino acid sequence
CC was deduced from a composite (see V27876) of isolated cDNA clones
CC 32 (see V27872) and TP2-15 (see V27875), obtained from a human
CC colon tumour cell line LIM1863 cDNA. Expressing TP2 in a cell is
CC used to increase telomerase activity and thus proliferation for
CC treatment of e.g. HIV infection, AIDS and ageing disorders, while
CC expressing an inactive mutant of TP2 (or molecule antisense to the
CC gene) is used to decrease telomerase activity, e.g. for treatment
CC of cancer. TP2 polypeptides can also be used to screen for agents
CC that inhibit TP2 activity or its binding to TRIP1 (see W61347) or
CC telomerase RNA, potentially useful therapeutically, also to raise
CC specific antibodies useful in immunoassays and therapeutically as
CC inhibitors. Also contemplated are transgenic animals in which the
CC TP2 gene has been inactivated or is overexpressed. TP2 polypeptides
CC are administered i.v., s.c. or orally, or they are delivered from
CC engineered cells or gene therapy vectors.
SQ Sequence 1154 AA;

```

Query Match 59.4%; Score 3505; DB 1; Length 1154;
 Best Local Similarity 62.4%; Pred. No. 0;
 Matches 719; Conservative 122; Mismatches 260; Indels 52; Gaps 13;

```

QY 1 MTRAPRCRAVRSLLSRVREVMPLATFVRLRGLPGRRLVQPGDKPIRYTLVAQCLVCMHW 60
Db 23 MTRAPRCRAVRSLLSRVREVMPLATFVRLRGLPGRRLVQPGDKPIRYTLVAQCLVCMHW 62
QY 61 GSQPPADLSHFQVSSKLKELVARVQRLCERNERNVLAFLGELLNEARGGPPMAFTSSVR 120
Db 83 DARPPAAPSPFQVSCLELVARVQRLCERNERNVLAFLGELLNEARGGPPMAFTSSVR 142
QY 121 SYLPNTVETLRVSGAMMLLSRVGDDLLVYLLAHCALYLLVPPSCAYOVCGSPLYOICA 180
Db 143 SYLPNTVETLRVSGAMMLLSRVGDDLLVYLLAHCALYLLVPPSCAYOVCGSPLYOICA 202
QY 181 TTDIWPVSASVYRTPVGRNFTNLRFLQIKSSRQEAQKPLALPSRGTKRHLSTSTS 240
Db 203 ATQARPPPHAS -GPRRLG-----CERAWNSVRAGVPLGLPAPGARRRGGSSASRS 253
QY 241 VPSAKKARCYVPVRVEGP-----HRQVLPSPGKSW-VPSPARSPVPTAEKDLSSK 292
Db 254 LPLKPRRARGAAPEPTPVGCGSWAHPRGRGSDRGFCVVSAPAR-----PAEATSLE 308
QY 293 GKVSLSLS-GSVCKKHPSTSLSPRONAFQLRP-FIETRHFLYSRGDGQERLNPSF 350
Db 309 GALSOTRSHSPSVGRQHAGPPTSRSRPRPDWTPCPVYAEATHFLYSSGD-KEQLRPSF 367
QY 351 LLSNLOPLNLTGARRLVEIFLGSRRPTSGPLCTRHLSSRYWQRPFLQOLLVNHAEQY 410
Db 368 LLSLRPSLTGARRLVETIFLGSRRPMPGTPRRLPLRQYQWQRPFLLELGNHACQY 427
QY 411 VRLLSHCRFTANQOQVTDAL-----NTSPHMLDLRLHSSPWQY 452
Db 428 GVLLKTHCPRAA---VTPAAGVAREKPGSVAAPEEDTPRRLVQLLRHSSPWQY 484

```

PT	and ageing	PT	and ageing
PS	Claim 2: Fig 2: 76pp: German.	PS	This sequence represents a novel human catalytic telomerase sub-unit
CC	This sequence represents a novel human catalytic telomerase sub-unit	CC	(hlc). This protein can be used in screening assays to identify
CC	modulators of telomerase and to treat or inhibit cellular disorders,	CC	death, defects and/or other pathological processes involving telomerase,
CC	particularly cancer and ageing (also suitable for this agents that	CC	stimulate, inhibit or mimic the activity of the subunit). Antisense
CC	nucleic acids inhibit telomerase action (by binding to specific mRNA),	CC	particularly in neoplastic cells and may be expressed in vivo. Antibodies
CC	and fragments of the protein, used as probes or primers, are used to	CC	diagnose telomerase-related conditions (especially neoplasia) by (i)
CC	detecting abnormal levels of the subunit protein in body fluids or	CC	tissues or (ii) by measuring the amount of the encoding nucleic acid.
CC	Expression of the nucleic acid encoding the subunit mRNA is confined to	CC	tumour cells, in contrast to the ubiquitous expression of the telomerase
CC	RNA subunit.	CC	RNA subunit.
CC	Sequence 1132 AA;	CC	Sequence 1132 AA;
QY	Query Match 59.4%; Score 3505; DB 1; Length 1132;	QY	Query Match 59.4%; Score 3505; DB 1; Length 1132;
DB	Best Local Similarity 62.4%; Pred. No. 0;	DB	Best Local Similarity 62.4%; Pred. No. 0;
DB	Matches 719; Conservative 122; Mismatches 260; Indels 52; Gaps 13;	DB	Matches 719; Conservative 122; Mismatches 260; Indels 52; Gaps 13;
QY	1 MTRAPRCVRSLLRSRYREVWPLATFVRRLGPEGRRLVQPGDKIYRTLVACQLVCMHW 60	QY	1 MTRAPRCVRSLLRSRYREVWPLATFVRRLGPEGRRLVQPGDKIYRTLVACQLVCMHW 60
DB	1 MPRAPRCVRSLLRSRYREVWPLATFVRRLGPEGRRLVQPGDKIYRTLVACQLVCMHW 60	DB	1 MPRAPRCVRSLLRSRYREVWPLATFVRRLGPEGRRLVQPGDKIYRTLVACQLVCMHW 60
QY	61 GSOPPPADLSFHQVSSLLKELVARVQRLCERNRNVLATGFEILLNEARGPPMAFTSSVR 120	QY	61 GSOPPPADLSFHQVSSLLKELVARVQRLCERNRNVLATGFEILLNEARGPPMAFTSSVR 120
DB	61 DARPPAAPSFQVSCLELVARVQRLCERNRNVLATGFEILLNEARGPPMAFTSSVR 120	DB	61 DARPPAAPSFQVSCLELVARVQRLCERNRNVLATGFEILLNEARGPPMAFTSSVR 120
QY	121 SYLNTVIEIRVSGANMLLSRVGDDLLVLLAHALYLLVPPSCAYQVCGSPLYOICA 180	QY	121 SYLNTVIEIRVSGANMLLSRVGDDLLVLLAHALYLLVPPSCAYQVCGSPLYOICA 180
DB	121 SYLNTVTDALRGSGANMLLSRVGDDLLVLLAHALYLLVPPSCAYQVCGSPLYOICA 180	DB	121 SYLNTVTDALRGSGANMLLSRVGDDLLVLLAHALYLLVPPSCAYQVCGSPLYOICA 180
QY	181 TTDIWPVSASRYRTPVGRNFTNLRFLOQIKSSROEAPKPLALSRGTRKRLSLTSTS 240	QY	181 TTDIWPVSASRYRTPVGRNFTNLRFLOQIKSSROEAPKPLALSRGTRKRLSLTSTS 240
DB	181 ATQARPPPHAS-GPRRLG-----CERAWNHSVREAGVPLGPAPGARRGGASRS 231	DB	181 ATQARPPPHAS-GPRRLG-----CERAWNHSVREAGVPLGPAPGARRGGASRS 231
QY	241 VPSAKKARCYVPVREEGP-----HRQVLPSPSGSW-VPSPARSPVPTAEKDLSSK 292	QY	241 VPSAKKARCYVPVREEGP-----HRQVLPSPSGSW-VPSPARSPVPTAEKDLSSK 292
DB	232 LFLPKRPRGGAAPFERTFVGQSWAHGPRTRGSDRGFCVVSFAR-----PAEATFLE 286	DB	232 LFLPKRPRGGAAPFERTFVGQSWAHGPRTRGSDRGFCVVSFAR-----PAEATFLE 286
QY	293 GRVSDLSL-GSVCKKHKPSSTSLSPRONAFOLRP-FIETRHFLSRGQGERLNPFS 350	QY	293 GRVSDLSL-GSVCKKHKPSSTSLSPRONAFOLRP-FIETRHFLSRGQGERLNPFS 350
DB	287 GALSGRHSHPVGRHAGPPSTSRPPRPMWDTCPVPIAETKHFYSSGD-KEQLRPSF 345	DB	287 GALSGRHSHPVGRHAGPPSTSRPPRPMWDTCPVPIAETKHFYSSGD-KEQLRPSF 345
QY	351 LLSNLOPNLTGARRLVEIFLGRSPTSGPLCRTHRLSRRYWQMRPLFOQLLVNHAECQY 410	QY	351 LLSNLOPNLTGARRLVEIFLGRSPTSGPLCRTHRLSRRYWQMRPLFOQLLVNHAECQY 410
DB	346 LLSLRPSLTGARRLVETIFLGRSPTSGPLCRTHRLSRRYWQMRPLFOQLLVNHAECQY 405	DB	346 LLSLRPSLTGARRLVETIFLGRSPTSGPLCRTHRLSRRYWQMRPLFOQLLVNHAECQY 405
QY	411 VRLSHCRFTANOQVTDAL-----NTSPHMLDLRLHSSPWQVY 452	QY	411 VRLSHCRFTANOQVTDAL-----NTSPHMLDLRLHSSPWQVY 452
DB	406 GVLLKTHCPRLAA---VTPAAGVCAREKPGQSVAAPEEDDPRLVQLLRHSSPWQVY 462	DB	406 GVLLKTHCPRLAA---VTPAAGVCAREKPGQSVAAPEEDDPRLVQLLRHSSPWQVY 462
QY	453 GFLRACLVKVSASLWGTNRHRRERFKNLKFISLGKYSKLSQLQELMKMKVEDCHWLSR 512	QY	453 GFLRACLVKVSASLWGTNRHRRERFKNLKFISLGKYSKLSQLQELMKMKVEDCHWLSR 512
DB	463 GFVRACLRLVPPGLWGSRRHRRERFKNLKFISLGKYSKLSQLQELMKMKVEDCHWLSR 522	DB	463 GFVRACLRLVPPGLWGSRRHRRERFKNLKFISLGKYSKLSQLQELMKMKVEDCHWLSR 522
QY	513 SPCKDRVPAEHRRLRERILATFLWMDTYVYVQLLSRFFYITESTFQXNRLFFYKRSWS 572	QY	513 SPCKDRVPAEHRRLRERILATFLWMDTYVYVQLLSRFFYITESTFQXNRLFFYKRSWS 572
DB	523 SPVGCVPAAEHLREILAKFLHMLSVYVVELLSRFFYITESTFQXNRLFFYKRSWS 582	DB	523 SPVGCVPAAEHLREILAKFLHMLSVYVVELLSRFFYITESTFQXNRLFFYKRSWS 582
QY	573 KLOSIGVROHLRRLRELRSQEVRRHQDTWLAMPICRLRFIPKPNGLRPIVNMYSMTG 632	QY	573 KLOSIGVROHLRRLRELRSQEVRRHQDTWLAMPICRLRFIPKPNGLRPIVNMYSMTG 632
DB	583 KLOSIGIRQLKRVQLRELSEAEVRRHQDPALTLRLRIPKPNGLRPIVNMYSMTG 642	DB	583 KLOSIGIRQLKRVQLRELSEAEVRRHQDPALTLRLRIPKPNGLRPIVNMYSMTG 642
QY	633 RALGRKQAOHFTORLKTFLSMLNRYERTKHPHLMGSSVLGMNDIYKTRAFVLRVRLDQ 692	QY	633 RALGRKQAOHFTORLKTFLSMLNRYERTKHPHLMGSSVLGMNDIYKTRAFVLRVRLDQ 692
DB	643 RTFRKRAELTSRVKALFSLVNLNERARRPGLGLGASVLGLDDIHRWRTFVLRVRAQDP 702	DB	643 RTFRKRAELTSRVKALFSLVNLNERARRPGLGLGASVLGLDDIHRWRTFVLRVRAQDP 702
QY	693 TPRMYFKADVGTGADYDAIPQGLKLVVAVNMIRHSESTYCIROYAVVRDSQGVHKSFR 752	QY	693 TPRMYFKADVGTGADYDAIPQGLKLVVAVNMIRHSESTYCIROYAVVRDSQGVHKSFR 752
DB	725 PPLEYFKVDVTDGAYDTIPQDRLETVIASIILK-PQNTYCVRYAVVQAAHGHVRKAFKS 783	DB	725 PPLEYFKVDVTDGAYDTIPQDRLETVIASIILK-PQNTYCVRYAVVQAAHGHVRKAFKS 783
QY	753 QVTLSLDQPYMGOFKHLQDSDASALRNSVVEIOSISMNESSSLDFDFLHFLRHSVVK 812	QY	753 QVTLSLDQPYMGOFKHLQDSDASALRNSVVEIOSISMNESSSLDFDFLHFLRHSVVK 812
DB	784 HVTSLDQPYMQFVAHLQET--SPLRDAVVEQSSSLNEASSGLDFVFLRMCCHAVR 841	DB	784 HVTSLDQPYMQFVAHLQET--SPLRDAVVEQSSSLNEASSGLDFVFLRMCCHAVR 841
QY	813 IGDRCYTQCGIPQGSLSLTLSCICFGDMENKLFAYVORDGLLRFVDDFLVTPHLDO 872	QY	813 IGDRCYTQCGIPQGSLSLTLSCICFGDMENKLFAYVORDGLLRFVDDFLVTPHLDO 872
DB	842 IRGKSVYQCGIPQGSLSLTLSCICFGDMENKLFAGIRDRDGLLRLVDDFLVTPHLTH 901	DB	842 IRGKSVYQCGIPQGSLSLTLSCICFGDMENKLFAGIRDRDGLLRLVDDFLVTPHLTH 901
QY	873 AKTFLSLVGVPEYCGMNLQTVNVPVPGTGLGAAPYQALPCHLPFWCGLLDQDT 932	QY	873 AKTFLSLVGVPEYCGMNLQTVNVPVPGTGLGAAPYQALPCHLPFWCGLLDQDT 932
DB	902 AKTFLSLVGVPEYCGMNLQTVNVPVPGTGLGAAPYQALPCHLPFWCGLLDQDT 961	DB	902 AKTFLSLVGVPEYCGMNLQTVNVPVPGTGLGAAPYQALPCHLPFWCGLLDQDT 961
QY	933 LEVFCDSYGAQTSIKTSLTFQSVFAGKTMRNKLLSVLRKCHGLFLDLQVNSLQTVCI 992	QY	933 LEVFCDSYGAQTSIKTSLTFQSVFAGKTMRNKLLSVLRKCHGLFLDLQVNSLQTVCI 992
DB	962 LEVQDSYSSYARTSIRASLTFRNGFKAGRNMRKLGVLRLKCHSLFLDLQVNSLQTVCT 1021	DB	962 LEVQDSYSSYARTSIRASLTFRNGFKAGRNMRKLGVLRLKCHSLFLDLQVNSLQTVCT 1021
QY	993 NIYKIFLLOAFRRHACVQLQPFQDORVKNLTFGLGISOASOCYAILKYKNPMTLKAS 1052	QY	993 NIYKIFLLOAFRRHACVQLQPFQDORVKNLTFGLGISOASOCYAILKYKNPMTLKAS 1052
DB	1022 NIYKILLQAFRRHACVQLQPFQDORVKNLTFGLGISOASOCYAILKYKNPMTLKAS 1081	DB	1022 NIYKILLQAFRRHACVQLQPFQDORVKNLTFGLGISOASOCYAILKYKNPMTLKAS 1081
QY	1053 GS---FPEPAHWTLCYQAFLLKLAHSVYIKCLLGLPRTAQKLLCKLPKATWTLKAAA 1109	QY	1053 GS---FPEPAHWTLCYQAFLLKLAHSVYIKCLLGLPRTAQKLLCKLPKATWTLKAAA 1109
DB	1082 GAAGPLPEAVQWLCHOAFLLKLRHRTVYVPLGLSLRTAQTLQSLRKLPTTLTLEAAA 1141	DB	1082 GAAGPLPEAVQWLCHOAFLLKLRHRTVYVPLGLSLRTAQTLQSLRKLPTTLTLEAAA 1141
QY	1110 DPALSTDFQTLID 1122	QY	1110 DPALSTDFQTLID 1122
DB	1142 NPALPSDFKTLID 1154	DB	1142 NPALPSDFKTLID 1154
RESULT	4	RESULT	4
W90251	W90251 standard; Protein; 1132 AA.	W90251	W90251 standard; Protein; 1132 AA.
ID	W90251; (first entry)	ID	W90251; (first entry)
AC	24-MAY-1999	AC	24-MAY-1999
DE	Human catalytic telomerase sub-unit protein.	DE	Human catalytic telomerase sub-unit protein.
KW	Human; catalytic telomerase subunit; therapy; diagnosis; htc; assay;	KW	Human; catalytic telomerase subunit; therapy; diagnosis; htc; assay;
KW	modulator; treatment; inhibit; cellular disorder; death; defect; cancer;	KW	modulator; treatment; inhibit; cellular disorder; death; defect; cancer;
KW	ageing; antisense; neoplastic cell; telomerase-related condition;	KW	ageing; antisense; neoplastic cell; telomerase-related condition;
KW	tumour cell.	KW	tumour cell.
OS	Homo sapiens.	OS	Homo sapiens.
PN	W09859040-A2.	PN	W09859040-A2.
PD	30-DEC-1998.	PD	30-DEC-1998.
PF	09-JUN-1998; E03468.	PF	09-JUN-1998; E03468.
PR	14-APR-1998; DE-016496.	PR	14-APR-1998; DE-016496.
PR	20-JUN-1997; DE-026329.	PR	20-JUN-1997; DE-026329.
PR	26-MAR-1998; DE-013274.	PR	26-MAR-1998; DE-013274.
PA	(FARB) BAYER AG.	PA	(FARB) BAYER AG.
PI	Hagen G, Siegmund H, Weichel W, Wick M, Zubov D;	PI	Hagen G, Siegmund H, Weichel W, Wick M, Zubov D;
DR	WPI; 99-081276/07.	DR	WPI; 99-081276/07.
DR	N-PSDB; V72117.	DR	N-PSDB; V72117.
PT	New catalytically active subunit of human telomerase - used in the	PT	New catalytically active subunit of human telomerase - used in the
PT	modulation of telomerase activity, particularly for treating cancer	PT	modulation of telomerase activity, particularly for treating cancer

Db 703 PPELYFVKVDTGAYDITPDRLETEVIAIHK-PQNTYCVRRYAVQAAHGHVRKAFKS 761
 Qy 753 QVTTSLDLPQYMGFLKHLQSDASALRNSVVIQSIISMNESSSLDFDLFLHLSHVVK 812
 Db 762 HVSTLTDLPQYMRQFVAHQET--SPLEDAVVIQSSSLNEASSGLDFVLFRECHHAVR 819
 Qy 813 IGRDCTQCQIGIPGSSLSLTLSCFGDMENKFLAEVQRDGLLRFVDDFLVTPHLQ 872
 Db 820 IRGKSYVOCQIGIPGSSLSLTLSCFGDMENKFLAEVQRDGLLRFVDDFLVTPHLTH 879
 Qy 873 AKTFLTLVHGVPYGYGMINLQKTVNFPVPGTLGGAAPQLPAHCLFPWCGLLDQT 932
 Db 880 AKTFLRTLVRGVPYGYGVNLRKTVNFPVDEALGATFQVMPAHGLFPWCGLLDTRT 939
 Qy 933 LEVFCDSYGAQTSIKTSLTFSQSVFKAGTKMNRKLLSVLRKLGHLDFLDQVNSLQVCI 992
 Db 940 LEVQSDYSYARTSIRASLTENRGEKAGNRMRKLFQVLRKCHSLFLDLQVNSLQVCI 999
 Qy 993 NIYKIFLQAYRHACVQLPFDQVRKNTLTFELGISSQASCCYAILKYNPGMTLKAS 1052
 Db 1000 NIYKILLQAYRHACVQLPFDQVQWKNPTFFLRVSDTASLCYSILKAKNAGMSLGAK 1059
 Qy 1053 GS---FPPEAAHLCYQAFLLKLAHSAVYKCLLGLPLRTAQKLLCRKLPEATWTLKAAA 1109
 Db 1060 GAAGPLPSEAVQWLCHQAFLLKLTTRHRTVYVPLLSLRTAQTLKRLPOTTLTALAAA 1119
 Qy 1110 DPALSTDFQITLD 1122
 Db 1120 NPALPSDFKTLTD 1132

RESULT 5

W47000
 ID W47000 standard; Protein; 1285 AA.
 AC W47000;
 DT 13-AUG-1998 (first entry)
 DE His tagged thiorodexin moiety and full length hTERT fusion protein.
 KW Human; telomerase reverse transcriptase; hTERT; TRT; diagnosis;
 OS Synthetic; cell proliferation; cancer; ageing; ribonucleoprotein.
 OS Homo sapiens.
 FH Key
 FT Location/Qualifiers
 FT Misc_difference 119..120
 FT /note= "enterokinase cleavage site"
 FT 120..1285
 FT /label= hTERT
 FT /note= "full length human telomerase reverse
 transcriptase"
 PN GB2317891-A.
 PD 08-APR-1998.
 PR 01-OCT-1997; 020890.
 PR 14-AUG-1997; US-915503.
 PR 01-OCT-1996; US-724643.
 PR 18-APR-1997; US-844419.
 PR 25-APR-1997; US-846017.
 PR 06-MAY-1997; US-851843.
 PR 09-MAY-1997; US-854050.
 PR 14-AUG-1997; US-911312.
 PR 14-AUG-1997; US-912951.
 PA (GERO-) GERON CORP.
 PA (UYTE-) UNIV TECHNOLOGY CORP.
 PI Andrews WH, Cech TR, Chapman KB, Harley C, Lingner J,
 PI Morin GB, Nakamura T, Harley CB;
 DR WPI; 98-171633/16.
 PT Pure and recombinant human Telomerase Reverse Transcriptase and its
 PT variants - are useful in the diagnosis, prognosis and treatment of
 PT cell proliferation conditions especially cancer and ageing
 PS Example 6; Page 223; 387pp; English.
 CC The present sequence represents a fusion protein from an example
 CC of the present invention which describes human telomerase reverse
 CC transcriptase (hTERT). The present invention also describes the
 CC following methods: (A) determining whether a test compound is

CC a modulator of hTERT, by detecting the change in hTERT recombinant
 CC protein or polynucleotide, on administration of the compound;
 CC (b) preparation of recombinant telomerase by contacting a protein
 CC preparation of hTERT with a telomerase RNA component; (C) detection
 CC of the hTERT RNA or protein in a sample by binding a relevant
 CC probe to the sample and detecting the complex formed or in the case of
 CC RNA detection, amplifying the product and correlating the presence of
 CC complex or amplification product with presence of hTERT in the sample;
 CC and (D) increasing the proliferation of a vertebrate cell by increasing
 CC hTERT expression; and (E) the use of an agent that causes an increase in
 CC cell vertebrate cell proliferation to create a medicament that inhibits
 CC ageing. A protein preparation of hTERT and the polynucleotide encoding
 CC hTERT can be used in the manufacture of medicaments for inhibiting the
 CC effect of ageing or cancer. Inhibitors of telomerase activity can be
 CC used to treat conditions that are associated with high telomerase
 CC activity. A protein preparation of hTERT can also be used in the new
 CC methods.
 SQ Sequence 1285 AA;

Query Match 59.3%; Score 3499; DB 1; Length 1285;
 Best Local Similarity 62.3%; Pred. No. 0;
 Matches 718; Conservative 122; Mismatches 261; Indels 52; Gaps 13;
 QY 1 MTRAPRCPAVRSLRLSRVREVMPLATFVRRRLGPEGRRLLVQPGDKIYRTLVAAQCLVCMHW 60
 Db 154 MPRAPRCRAVRSLRLSRHYREVLPATFVRRRLGPOGWRVLRQGDPAARFALVAQCLVCPW 213
 QY 61 GSQPPADLSFHQVSSSLKELVARVQVRLCERNERNVLAFFGELLNEARGPPMAFTSSVR 120
 Db 214 DARPPAAPSPFQVSCLELVARVLRQRCAGKAVLAFGEALLDCARGPPEAFTTSVR 273
 QY 121 SYLPTVETIETRVSGAWMLLSRVGDDLLVLLAHACALYLLVPPSCAYQVCGSPLYQICA 180
 Db 274 SYLPTVTDALRGSGAGLLLRVGDVLLHLLARCALVFLVAPSCAYQVCGPPLYQIGA 333
 QY 181 TTDIWPVSASVPRPRVGRNFTNRLFLQIKSSSRQEPAPKPLAPSRGTRKHLSTSTS 240
 Db 334 ATQARPPPHAS-GPRRLG-----CERAWNSVREAGVPLGLPAPGARRRGSASRS 384
 QY 241 VPSAKKARCPVPRVEEGP-----HRQVLPSPGKSW-VPSPARSDEVTAKDLSSK 292
 Db 385 LPLKPRPRGAPEPPTPVGGGWAHPGRTGRDPSDFCVVSPAR-----PABEATSLE 439
 QY 293 GKVSLSLS-GSVCKKHKPSSTSLSPRONAFOLRP-FIETRHFLYSGDQERLNPSPF 350
 Db 440 GALSOTRSHSPVGRQHHAGPPTSRRPWPDTCPVPVYAEKHFLLYSSGD-KCOLRPSF 498
 QY 351 LLSNLPNLTCARLVEIIFLGSRRPTSGPLCRTHRSRRYWMRPLFQQLLVNHAECQY 410
 Db 499 LLSLRPSLTGARLVEITFLGSRPMPGTPLRLPRLPQRYWQMRPLFLELLGNHACQY 558
 QY 411 VRLLRSHCRFTANQOVTDAL-----NTSPHMLDLRLHSSPQVY 452
 Db 559 GVLLKTHCPLEAA---VTPAAGVCAREKPGSVAAPEEDTPRLVQLLRQHSPPQVY 615
 QY 453 GFLRACLVVSVASLWGTNRHNERFFNKLKFIISLGRYKLSLQELMKWKVDECHWLS 512
 Db 616 GFVRACLRLVPPGLWGSRRHNERFLNKKFISLGHAKLSLQELTWKSVRDCALRR 675
 QY 513 SPGKDRVPAAEHRLRERILATFLEWMDTVVQLLRSFFYITESTFQKNLFFYRKSVWS 572
 Db 676 SPGVGCPAAEHLRLEILAKFLHMLSVVVELLSFFYVTTETTFQKNLFFYRKSVWS 735
 QY 573 KLSIGVYRQHLRVRRLRELSQEEVRRHODTWLAMPICRLRFIPKPNGLRIPVNNSTSMGT 632
 Db 736 KLSIGIRQHLKRVQLRELSAEVRRHARPALLTSRLRFIPAPDGLRIPVNDYVYGA 795
 QY 633 RALGRKQAOHQFTQRLKTLFSLMYNERTKPHLMGSSVGLGMDIYRTWRAFLVRALDQ 692
 Db 796 RTFRREKRAERLTSRVKALFSVLNRYERARRPGLLGASVLGLDDIHRARWTFVLVRAQDP 855
 QY 693 TPRMYFVKVDGTGAYDAIPQGLVEVVANNIRHSESTYCIQRYAVVRRDSQGVHKSFR 752

856 PPELIFVKVDVTGAYDITIPDRLTEVIASLIK-PONTYCVRRYAVVQKAAGHVRKAFKS 914
 QY 753 QVTTLSLQPMQOFLKHLQSDASALRNSVIEQISMNESSESSDFDLFLHFLRHSVVK 812
 Db 915 HVSTLTLQPMQOFLKHLQSDASALRNSVIEQISMNESSESSDFDLFLHFLRHSVVK 972
 QY 813 IGDRCTOCQOIGOGSSLLSLCSLFCGDMENKLFQVQSDGGLLRFVDDFLAVTPLHQ 872
 Db 973 IRGKSYQCQOIGOGSSLLSLCSLFCGDMENKLFQVQSDGGLLRFVDDFLAVTPLHQ 1032
 QY 873 AKTFLSLVHGVPYEGCMINLQKTVNFPVPGTGLGAAPYQLPALHGLFPCWGLLDTQT 932
 Db 1033 AKTFLSLVHGVPYEGCMINLQKTVNFPVPGTGLGAAPYQLPALHGLFPCWGLLDTQT 1092
 QY 933 LEVFCDSYGAQTSIKTSLTFQSVFKAGTKMRNKLKLVRLKCHGLFLDLQVNSLQTVCI 992
 Db 1093 LEVFCDSYGAQTSIKTSLTFQSVFKAGTKMRNKLKLVRLKCHGLFLDLQVNSLQTVCI 1152
 QY 993 NIYKIFLQAVRHACVIOQLPFDQVRKNTFFLGIITSSOASCCYAILKVNKPGMTLKAS 1052
 Db 1153 NIYKIFLQAVRHACVIOQLPFDQVRKNTFFLGIITSSOASCCYAILKVNKPGMTLKAS 1212
 QY 1053 GS---FPPEAAHLCYQAFLLKLAHVIYKCLLGLPLTAQKLCRLKPLPMTILKAAA 1109
 Db 1213 GAAGPLPSEAVOWLCHQAFLLKLAHVIYKCLLGLPLTAQKLCRLKPLPMTILKAAA 1272
 QY 1110 DPALSTDFQTLTD 1122
 Db 1273 NPALPSPDFKTLTD 1285
 RESULT 6
 W1376
 ID W1376 standard; Protein; 1132 AA.
 AC W1376;
 DE Human telomerase catalytic subunit referred to as hstr2.
 KW Catalytic subunit; human; telomerase; telomere maintenance;
 diagnosis; treatment; cancer.
 OS Homo sapiens.
 PN W09837181-A2.
 PD 27-AUG-1998.
 PR 20-FEB-1998; U03404.
 PR 30-OCT-1997; US-064322.
 PR 20-FEB-1997; US-038750.
 PR 20-MAY-1997; US-047151.
 PR 01-AUG-1997; US-054549.
 PR 14-AUG-1997; US-055762.
 PA (WHED) WHITEHEAD INST BIOMEDICAL RES.
 PI Counter CM, Meyerson M, Weinberg RA;
 DR WPI: 98-495367/42.
 DR N-PSDB: V60320.
 PT New isolated human telomerase catalytic sub-unit gene - used to
 develop products for increasing or reducing the life span of cells
 such as cancer cells or transformed cells
 PS Claim 5; Fig 6; 96pp; English.
 CC The present sequence represents the catalytic subunit of a human
 telomerase holoenzyme. Disruption of the telomerase gene alters
 telomere maintenance. The DNA is essential for telomerase activity,
 and the protein is physically associated with telomerase and a
 constituent of active telomerase complex. The products can be used
 for increasing or reducing the lifespan of cells such as cancer cells
 or transformed cells. They can also be used in the diagnosis and
 treatment of malignancies. In addition, cells with a longer lifespan
 can be transplanted into or grafted onto an individual (e.g. as skin
 grafts, as systems for delivery of therapeutic proteins, such as hormones
 and enzymes), to whom they provide therapeutic benefit.
 SQ Sequence 1132 AA;

Query Match 59.3%; Score 3498; DB 1; Length 1132;
 Best Local Similarity 62.3%; Pred. No. 0;

Matches 718; Conservative 122; Mismatches 261; Indels 52; Gaps 13;
 QY 1 MTRAPRCVAVRSLRSRYREVWPLATFVRRLLGPEGRRLVQCPDKIYRTLVQAQCLVCMHW 60
 Db 1 MTRAPRCVAVRSLRSRYREVWPLATFVRRLLGPEGRRLVQCPDKIYRTLVQAQCLVCMHW 60
 QY 61 GSQPPADLSHFQVSSSLKELVARVQRCERNERNVLAFFGELLNEARGGPPMAFTSSVR 120
 Db 61 DARPPPAAPSPFQVSSSLKELVARVQRCERNERNVLAFFGELLNEARGGPPMAFTSSVR 120
 QY 121 SYLPTNVTIETLRSVSAWMLLSRVGDDLLVLAHALYLLVPPSCAYQVCSPLIYQICA 180
 Db 121 SYLPTNVTIETLRSVSAWMLLSRVGDDLLVLAHALYLLVPPSCAYQVCSPLIYQICA 180
 QY 181 TTDIWPSVSAVYRTPVGRNFTNLRFLLQIKSSRSQCAPKPLALPSRGTKRHLSLTSTS 240
 Db 181 ATQARPPPHAS-GPRRLUG-----CERAWNHSVREAGVPLGLPAPGARRRGSSASRS 231
 QY 241 VPSAKKARCYVPRVEEGP-----HROVLTPSGKSW-VSPARSPEVPTAEKDLSSK 292
 Db 241 VPSAKKARCYVPRVEEGP-----HROVLTPSGKSW-VSPARSPEVPTAEKDLSSK 292
 QY 293 GKYSDDLSS-GSVCKKHKPSSTLSLSPRQNAFQLRP-FIETRHFLYSRGDQOERLNSF 350
 Db 293 GKYSDDLSS-GSVCKKHKPSSTLSLSPRQNAFQLRP-FIETRHFLYSRGDQOERLNSF 350
 QY 351 LLSNLPNTGARRUVEIIFLGSRRPTSGPLCTHRLSRRYQWMPRLFOQLLVNHAECY 410
 Db 351 LLSNLPNTGARRUVEIIFLGSRRPTSGPLCTHRLSRRYQWMPRLFOQLLVNHAECY 410
 QY 411 VRLLSHCRFTANQOVTDAL-----NTSPHMLDMLLRHSSPWQY 452
 Db 411 VRLLSHCRFTANQOVTDAL-----NTSPHMLDMLLRHSSPWQY 452
 QY 453 GFLRACICKVYVSASLWGTNRHNERFFKNLKFISLQYKLSLQELQELWKMVKVDCWHLRS 512
 Db 453 GFLRACICKVYVSASLWGTNRHNERFFKNLKFISLQYKLSLQELQELWKMVKVDCWHLRS 512
 QY 513 SPKDRVPAAEHLRLERILATFLFWMVLLSFFVITESTFQKRLFFYRSVWS 572
 Db 513 SPKDRVPAAEHLRLERILATFLFWMVLLSFFVITESTFQKRLFFYRSVWS 572
 QY 573 KLOSIGVROHLERVLRELSEVRRHQDTWLAIPICRLRFIPKPNGLRPVNNYSMTG 632
 Db 573 KLOSIGVROHLERVLRELSEVRRHQDTWLAIPICRLRFIPKPNGLRPVNNYSMTG 632
 QY 632 RALGRRKQAOHQFTORLKTFLSMLNYERTKPHLMGSLVGMNDIYRTWRAFLVRALDQ 692
 Db 632 RALGRRKQAOHQFTORLKTFLSMLNYERTKPHLMGSLVGMNDIYRTWRAFLVRALDQ 692
 QY 692 RTFRERKRAERLTSVKALFSVLNYERARRPGLIGASVLGLDDIHRARWTFVLVRAQDP 702
 Db 692 RTFRERKRAERLTSVKALFSVLNYERARRPGLIGASVLGLDDIHRARWTFVLVRAQDP 702
 QY 702 TPRMYFVKADVTGAYDAIPOGKLEVVVANNIRHSESTYCIQYAVVRRDSQGVHKSFR 752
 Db 702 TPRMYFVKADVTGAYDAIPOGKLEVVVANNIRHSESTYCIQYAVVRRDSQGVHKSFR 752
 QY 752 PPELIFVKVDVTGAYDITIPDRLTEVIASLIK-PONTYCVRRYAVVQKAAGHVRKAFKS 761
 Db 752 PPELIFVKVDVTGAYDITIPDRLTEVIASLIK-PONTYCVRRYAVVQKAAGHVRKAFKS 761
 QY 761 QVTTLSLQPMQOFLKHLQSDASALRNSVIEQISMNESSESSDFDLFLHFLRHSVVK 812
 Db 761 QVTTLSLQPMQOFLKHLQSDASALRNSVIEQISMNESSESSDFDLFLHFLRHSVVK 812
 QY 812 HVSTLTLQPMQOFLKHLQSDASALRNSVIEQISMNESSESSDFDLFLHFLRHSVVK 872
 Db 812 HVSTLTLQPMQOFLKHLQSDASALRNSVIEQISMNESSESSDFDLFLHFLRHSVVK 872
 QY 872 AKTFLSLVHGVPYEGCMINLQKTVNFPVPGTGLGAAPYQLPALHGLFPCWGLLDTQT 932
 Db 872 AKTFLSLVHGVPYEGCMINLQKTVNFPVPGTGLGAAPYQLPALHGLFPCWGLLDTQT 932
 QY 933 LEVFCDSYGAQTSIKTSLTFQSVFKAGTKMRNKLKLVRLKCHGLFLDLQVNSLQTVCI 992
 Db 933 LEVFCDSYGAQTSIKTSLTFQSVFKAGTKMRNKLKLVRLKCHGLFLDLQVNSLQTVCI 992
 QY 993 NIYKIFLQAVRHACVIOQLPFDQVRKNTFFLGIITSSOASCCYAILKVNKPGMTLKAS 1052
 Db 993 NIYKIFLQAVRHACVIOQLPFDQVRKNTFFLGIITSSOASCCYAILKVNKPGMTLKAS 1052
 QY 1000 NIYKIFLQAVRHACVIOQLPFDQVRKNTFFLGIITSSOASCCYAILKVNKPGMTLKAS 1059
 Db 1000 NIYKIFLQAVRHACVIOQLPFDQVRKNTFFLGIITSSOASCCYAILKVNKPGMTLKAS 1059

ID W56101 standard; Protein; 1405 AA.
AC W56101;
DT 13-AUG-1998 (first entry)
DE Enhanced green fluorescent protein and hTERT fusion protein.
KW Human; telomerase reverse transcriptase; hTERT; RT; diagnosis;
KW prognosis; cell proliferation; cancer; ageing; ribonucleoprotein.
OS Synthetic.
OS Homo sapiens.
FH Key
FT Location/Qualifiers
FT 1..250
FT /note= "enhanced green fluorescent protein fragment"
FT 276..1405
FT /note= "hTERT protein fragment"
PD GB2317891-A.
PN 08-APR-1998.
PP 01-OCT-1997; 020890.
PP 14-AUG-1997; US-915503.
PK 01-OCT-1996; US-724643.
PK 18-APR-1997; US-844419.
PR 23-APR-1997; US-846017.
PR 06-MAY-1997; US-851843.
PR 09-MAY-1997; US-854050.
PR 14-AUG-1997; US-911312.
PR 14-AUG-1997; US-912951.
PA (GERO-) GERON CORP.
PA (UYTE-) UNIV TECHNOLOGY CORP.
PI Andrews WH, Cech TR, Chapman KB, Harley C, Lingner J,
PI Morin GB, Nakamura T, Harley CB;
DR WPI: 98-171633/16.
PT Pure and recombinant human Telomerase Reverse Transcriptase and its
PT variants - are useful in the diagnosis, prognosis and treatment of
PT cell proliferation conditions especially cancer and ageing
PS Example 15: Page 269-270; 387pp; English.
CC The present sequence represents a fusion protein from an example
CC of the present invention which describes human telomerase reverse
CC transcriptase (hTERT). The present invention also describes the
CC following methods: (A) determining whether a test compound is
CC a modulator of hTERT, by detecting the change in hTERT recombinant
CC protein or polynucleotide, on administration of the compound;
CC (B) preparation of recombinant telomerase by contacting a protein
CC of the hTERT RNA or protein in a sample by binding a relevant
CC RNA detection, amplifying the product and correlating the presence of
CC and (D) increasing the proliferation of a vertebrate cell by increasing
CC hTERT expression; and (E) the use of an agent that causes an increase in
CC cell vertebrate cell proliferation to create a medicament that inhibits
CC ageing. A protein preparation of hTERT and the polynucleotide encoding
CC effect of ageing or cancer. Inhibitors of telomerase activity can be
CC used to treat conditions that are associated with high telomerase
CC activity. A protein preparation of hTERT can also be used in the new
CC methods.
SQ Sequence 1405 AA;

Query Match 58.7%; Score 3463; DB 1; Length 1405;
Best Local Similarity 61.9%; Pred. No. 0;
Matches 713; Conservative 121; Mismatches 266; Indels 52; Gaps 13;

QY 1 MTRAPCAPVRLSLRSRYREVWPLATFVRLGPGGRRLVQDPKPIRLVLAQCLVCMHW 60
Db 276 MPRAPCRVRLSLRSRYREVWPLATFVRLGPGGRRLVQDPKPIRLVLAQCLVCMHW 60
QY 61 GSOPPPADLSFHOVSSLSKELVARVORLCERNERNYLAFCFELLNEARGPPMAFTSSVR 120
Db 336 DARPPPAAPSFQVSCLELVARVORLCERNERNYLAFCFELLNEARGPPMAFTSSVR 120
QY 121 SYLPTNTVETLRVSGAWMLLSRVGDDLLVLLAHCALVLLPPSCAYOVCGPLYQICA 180
Db 396 SYLPTNTVETLRVSGAWMLLSRVGDDLLVLLAHCALVLLPPSCAYOVCGPLYQICA 455

QY 181 TTDIWPVSYSASYRTPRPVGRNFTNLRFLOQIKSSROEAPKPLALPSRGTKRHLSTLS 240
Db 456 ATOARPPPHAS-GPRRLG-----CERAWNHSVREAGVPLGLPAGARRGGSSASRS 506
QY 241 VPSAKKARCYPVPRVEEGP-----HROVLPPTSGKSW-VPSPARSPVEPTAEKDLSSK 292
Db 507 LPLPKRPRGAAPPERTPVGGQSWAHPGRTGRDPSDRGFCVSPAR-----PAEATSLE 561
QY 293 GKVSLSLS-GSVCKKHKPSSTLSLSPQNAFOLRFETRIHFLYLRSGQGERLPSFL 351
Db 562 GALSCTRHSHPSVGRHQHAGPPSTSRPPRWDTPCPVYAEATRFYSSSGD-KEQLRPSFL 630
QY 352 LSNLQPNLTGARRLVEIFLGSRPRTSGPLCRTHRLSRRYQWQMRPLFQOQLVNHACQYV 411
Db 621 LSSLRPSLTGARRLVETIFLGSRPMPGTPRLPLPQRYQWQMRPLFLELLGNHACQYV 411
QY 412 RLLRSHCFRTANQOVTDAL-----NTSPHMLDMLRLHSSSPWQYVG 453
Db 681 VLLKTHCPLRAA---VTPAAGVCAREKPGQSVAAPEEDTDPRLVOLLROHSSPWQYVG 737
QY 454 FLRACICKVYSASLMGTRHNERFFKNLKKFISLGRYKGLSLQELMKMKVEDCHWLRS 513
Db 738 FVRACLRRLVPPGLWGSRRHNERFLNRTKKFISLGRHAKLSLQELTKWMSVRCDAWLRS 797
QY 514 PGKDRVPAAEHRLRLERILATELFWLMDTVVQLLSFFYITESTFOKNLFFYRKSVMK 573
Db 798 PGVGCVPAAEHLREILAKFLHMLMSVYVVELLSFFYITESTFOKNLFFYRKSVMK 857
QY 574 LQSIGVROHLERVLRELSEAEVROHREARPAALLTSRLRFIPKPGCLRPVNMDFVVGAR 917
Db 858 LQSIGVROHLERVLRELSEAEVROHREARPAALLTSRLRFIPKPGCLRPVNMDFVVGAR 917
QY 634 ALGRRQAQOHTORLKTLESMNLNYERTKHPHMGSSVIGMNDIYRTWAFVLRVLRDQT 693
Db 918 TFRERARERLTSRVKALEFSLNYERARPGGLGASVLGLDDIHRARTFVLRVLRDQT 693
QY 694 PRMYFVKADVTGAYDAIPQKGLVEVANNIRHSESTYCIROYAVVVRDSDOQVHKSFRQ 753
Db 978 PELYFVKVDTGAYDTIPQDRLEVIASIK-PQNTYCVRYAVVQKAAHGVKAFKSH 1036
QY 754 VTTLSLQPYMGQFLLKHLQSDASALRNSVIEOSISMNESSSLDFDFLHFLRHSVVKI 813
Db 1037 VSTLDLQPYMHQFVAHQET--SPLRDVYIEQSSSLNEASSGLDFVFLRFMCHHAVRI 1094
QY 814 GDRCTQCGIPQGSSTLCLSCFCGDMENKLFAYVORDGLLIRFVDDFLVTPHLDQA 873
Db 1095 RGSYVQCGIPQGSSTLCLSCFCGDMENKLFAYVORDGLLIRFVDDFLVTPHLDQA 873
QY 874 KTFSLTVHGVPEYGCMLNOKTVVNPFPVPGTLGGAAPYQLPARCLFPWCGLLDFTOL 933
Db 1155 KTFSLTVHGVPEYGCMLNOKTVVNPFPVPGTLGGAAPYQLPARCLFPWCGLLDFTOL 933
QY 934 EYFCDYSYAOITSIKTFSQVFKAGTKRNLKLSVLRLKCHGLFLDLQVNSLQTVGIN 993
Db 1215 EVQSDYSYARTSRASVTFNRGFRGNRRKLFGLVRLKCHSLFLDLQVNSLQTVGIN 1274
QY 994 IYKFLQAVYFHACVITQLPDQVRKRLTFFLGIISQASCCVAILKYKNPGMTLKASG 1053
Db 1275 IYKFLQAVYFHACVITQLPDQVRKRLTFFLGIISQASCCVAILKYKNPGMTLKASG 1053
QY 1054 S---FPPEAAHLCYQAFLLKLAHSHYIKCLLGPRTAQKLCRKLPEATMTILKAAAD 1110
Db 1335 AAGPLPSEAVQWCHQAFLLKLRHRTVYVPLGLSL-TAQTLRSRKLPGTTLTALEAAN 1393
QY 1111 PALSTDFQITLD 1122
Db 1394 PALPSDFKITLD 1405

RESULT 10
W47006
ID W47006 standard; Protein; 1150 AA.
AC W47006;

Mon Aug 16 10:21:26 1999

DT	13-AUG-1998 (first entry)	293	GVSDLSLS-GSVCKHKPSTSLSPRQAFQRLP-FIETRHLYSRGQDQRLNPSF	350
DE	Glutathione-S-transferase and hTERT fusion protein 6.			
DE	Human; telomerase reverse transcriptase; hTERT; TRT; diagnosis;			
KW	prognosis; cell proliferation; cancer; ageing; ribonucleoprotein.			
OS	Synthetic.			
OS	Homo sapiens.			
PN	GB2317891-A.			
PD	08-APR-1998.			
PF	01-OCT-1997; 020890.			
PR	14-AUG-1997; US-915503.			
PR	01-OCT-1996; US-724643.			
PR	18-APR-1997; US-844419.			
PR	25-APR-1997; US-846017.			
PR	06-MAY-1997; US-851843.			
PR	09-MAY-1997; US-854050.			
PR	14-AUG-1997; US-911312.			
PR	14-AUG-1997; US-912951.			
PA	(GERO-) GERON CORP.			
PI	(UYTE-) UNIV TECHNOLOGY CORP.			
PI	Andrews WH, Cech TR, Chapman KB, Harley C, Lingner J,			
PI	Morin GB, Nakamura T, Harley CB;			
DR	WPI; 98-171633/16.			
PT	Pure and recombinant human Telomerase Reverse Transcriptase and its			
PT	variants - are useful in the diagnosis, prognosis and treatment of			
PT	cell proliferation conditions especially cancer and ageing			
PT	Example 6; Page 231-232; 387pp; English.			
PS	The present sequence represents a fusion protein from an example			
CC	of the present invention which describes human telomerase reverse			
CC	transcriptase (hTERT). The present invention also describes the			
CC	following methods: (A) determining whether a test compound is			
CC	a modulator of hTERT, by detecting the change in hTERT recombinant			
CC	protein or polynucleotide, on administration of the compound;			
CC	(B) preparation of recombinant telomerase by contacting a protein			
CC	preparation of hTERT with a telomerase RNA component; (C) detection			
CC	of the hTERT RNA or protein in a sample by binding a relevant			
CC	probe to the sample and detecting the complex formed or in the case of			
CC	RNA detection, amplifying the product with presence of hTERT in the sample;			
CC	complex or amplification product with presence of hTERT in the sample;			
CC	and (D) increasing the proliferation of a vertebrate cell by increase in			
CC	hTERT expression; and (E) the use of an agent that causes an increase in			
CC	cell vertebrate cell proliferation to create a medicament that inhibits			
CC	ageing. A protein preparation of hTERT and the polynucleotide encoding			
CC	hTERT can be used in the manufacture of medicaments for inhibiting the			
CC	effect of ageing or cancer. Inhibitors of telomerase activity can be			
CC	used to treat conditions that are associated with high telomerase			
CC	activity. A protein preparation of hTERT can also be used in the new			
CC	methods.			
SQ	Sequence 1150 AA;			
	Query Match 56.3%; Score 3320; DB 1; Length 1150;			
	Best Local Similarity 61.2%; Pred. No. 7.8e-299;			
	Matches 708; Conservative 122; Mismatches 263; Indels 64; Gaps 23;			
QY	1 MTRAPCPAVRSLRSREWVPLATFVRRLGPEGRRLLVQDPKRYITLVAQCLVCMHW 60			
DB	1 MPRAPCRVAVRSLRSREWVPLATFVRRLGPEGRRLLVQDPKRYITLVAQCLVCMHW 60			
QY	61 GSOPPPADLSFHOVSKELVARVQRLCERNERNVLAFFGELLNEARGGPPMAFTSSVR 120			
DB	61 DARPPAPAFRQVSCLEKELVARVQRLCERNERNVLAFFGELLNEARGGPPMAFTSSVR 120			
QY	121 SYLPTNTVIETLVSGAWMLLSRVGDDLLVYLALHACALVLLVPPSCAYQVCGSPLYQICA 180			
DB	121 SYLPTNTVIETLVSGAWMLLSRVGDDLLVYLALHACALVLLVPPSCAYQVCGSPLYQICA 180			
QY	181 TTDWPSVASRYRTPVGRNFTNLRLQOIKSSSRQEAAPKPLALPSRGTKHLSTSTS 240			
DB	181 ATQARPPPHAS-GPRRLG-----CERAWNHSVREAGVPLGLPAPGARRRGGSAGRS 231			
QY	241 VPSAKKARCVPRVREBGP-----HRQVLPTPSGKSW-VPSAPSPVPTAEKDLSSK 292			
DB	232 LPLPKRPRRGAAPERTPTVPGQGSWAHPGTRGDSRGFCVSPAR-----PAEATSLSE 286			
QY	293 GKVSDLSLS-GSVCKHKPSTSLSPRQAFQRLP-FIETRHLYSRGQDQRLNPSF 350			
DB	287 GALSCTRHSHPVGRHQHAGPSTSRPPRWDTPCPPVYATKHFLLYSSGD-KQLRPSF 345			
QY	351 LLSNLQPNLTGARRLVEIIFIGSRPRTSGPICRTHLSRRYQWMPFLFOQLLVNHAECQY 410			
DB	346 LLSLRLPSTLGARRLVETIFLGS-PWMEGTPRRLPRLPQRYQWMPFLFELLGNAQCPY 404			
QY	411 VLLRSHCFRTANQOVTDAL-----NTSPPLMDLLRLHSSPWQVY 452			
DB	405 GVLKTHCPLRAA---VTPAAGVCAKREKPGQSVAAPEEEDTDPRLVOLLRQHSQWQVY 461			
QY	453 GFLRACILCKVVSASL-WGTRHNERFFKMLKPFISLGKYGKLSQLQELMKMKVECHWLRL 511			
DB	462 GFVRACLRLLVPPGLEVGRHNERFLRNTKRFISLGKHAQLSLOELATWKMSVROCAMLR 521			
QY	512 SSPGKDRVPAAEHRLRERILATFLWMDTYVOLLRSFFYITESTFQKNRLFFYKSWY 571			
DB	522 RSPGVGCVPAAEHRLRERILAKFLHLMMSVYVVELLSFEW-TEITFQKNRLFFYKSWY 580			
QY	572 SKLQSIGVROHLRERLRELSQEEVRHQDTWLAMPICRLRIFPKPGLRPIVNMYSYMG 631			
DB	581 SKLQSIGVROHLRERLRELSQEEVRHQDTWLAMPICRLRIFPKPGLRPIVNMYSYMG 640			
QY	632 TRALGRKQAHQHTQRLKTLFSLMNYERTKHPHLMGSSVLMGNDIYRTWRAFLVRLRALD 691			
DB	641 ARTRERKRAERLTSRYKALFVNLNERARRPGLLGASVLGDDIHRAMRTFVLVRYAQQD 700			
QY	692 QTPRMVFKADVTGAYDAIPQGLVEYVANNIRHSESTYCIQIYAVYVRDSQOVHKSFR 751			
DB	701 PPPELYFKVDVTGAYDTIIPQDRLETVIASLIK-PQNTYCVRRYAVYVQKAAHGHVRKAFK 759			
QY	752 ROVTLSLDQPMYQFQLKHLQDSASALRNSVYEQSISM-NESSSLDFDFHFLHRSYV 810			
DB	760 SHVSTLTLDQPMYQFVAHQDET--SPLRDAVYEQSSSLETEASSGLDFVFLRFMCHHA 817			
QY	811 VKIGDRCTQCGIPQGSSTLSCFCFGDMENKLFQVORDGLLRLFRVDDFLLVTPHL 870			
DB	818 VERGKS--WQCGIPQGSSTLSCFCFGDMENKLFQVORDGLLRLFRVDDFLLVTPHL 875			
QY	871 DOAKTFTLTVHGVPEYCGMNLQKTVVNFVEGTGGGAAPYOLPAHCLFPW-CGLILD 929			
DB	876 THAKTFTIRLVG-PEYCGVNLKTV--FPVEDALGGTAFOVMPAHGLFPWVCGLLLD 932			
QY	930 TQTFVCDYSYATQTSIKTSLTQSFQVFKAG-TMRNKLKLSVLRLKCHGLFLDQVNSLQ 988			
DB	933 TRILEVQSDSSYATQTSIKTSLTQSFQVFKAG-TMRNKLKLSVLRLKCHGLFLDQVNSLQ 992			
QY	989 TVCINIKIFLLOAYRHACVQLPFDQVRKNLTFGLIILSSQASCCYAILKVKNPQMT 1048			
DB	993 TVCTNIYKILLQAYRHACVQLPFDQVRKNLTFGLIILSSQASCCYAILKVKNPQMT 1052			
QY	1049 LRAGS---FPPEAAHLCYQAFLLKLAHSHYIYKLLGPLRTAOKLCKRLKLPENATMIL 1105			
DB	1053 LGAKGAAGPLPSEAVOWLCHQAFLLKLTFRHYTYVPLLGSLRTAQTQLSRKLPGTTLTAL 1112			
QY	1106 KAAADPALSTDFQITLD 1122			
DB	1113 EAAANPALS-DFKTILD 1128			
RESULT	11			
W61349	standard; Protein; 949 AA.			
ID	W61349;			
AC	12-OCT-1998 (first entry)			
DE	Human telomerase protein 2 (TP2) partial polypeptide.			
KW	TP2; human; telomerase protein 2; cancer; AIDS; ageing; therapy.			
OS	Homo sapiens.			
OS	Location/Qualifiers			
FT	640..940			
FT	/note= "Claim 22"			

FT	Protein	
FT	1..563	/note="Claim 24"
FT	Protein	
FT	1..640	/note="Claim 24"
FT	Protein	
FT	696..940	/note="Claim 24"
FT	Protein	
FT	696..953	/note="Claim 24"
FT	Region	
FT	582..587	/note="Claim 24"
FT	Region	
FT	644..648	
FT	Region	
FT	852..857	
FT	Region	
FT	884..8894	
PN	WO9821343-A1.	
PD	22-MAY-1998.	
PF	13-NOV-1997; U21248.	
PR	16-OCT-1997; US-951733.	
PR	15-NOV-1996; US-871189.	
PR	11-JUN-1997; US-873039.	
PA	(AMGE-) AMGEN CANADA INC.	
PA	(AMGE-) AMGEN INC.	
PI	Harrington LA, Robinson MO;	
PI	WPI; 98-297946/26.	
DR	N-PSDB; V27872.	
PT	New nucleic acid encoding human telomerase protein-2 - used for	
PT	regulating telomerase activity, e.g. for treating cancer or acquired	
PT	immune deficiency syndrome	
PS	Claim 1d; Fig 6; 150pp; English.	
CC	This polypeptide comprises a large portion of human telomerase	
CC	protein 2 (TP2), a novel protein from partial cDNA clone 32 (see	
CC	amino acid sequence was deduced from partial cDNA clone 32 (see	
CC	V27872). A full-length polypeptide sequence (see W61350) is also disclosed.	
CC	Expressing TP2 in a cell is used to increase telomerase activity and	
CC	thus proliferation for treatment of e.g. HIV infection, AIDS and	
CC	aging disorders, while expressing an inactive mutant of TP2 (or	
CC	molecule antisense to the gene) is used to decrease telomerase	
CC	activity, e.g. for treatment of cancer. TP2 polypeptides can also	
CC	be used to screen for agents that inhibit TP2 activity or its	
CC	binding to TRIP1 (see W61347) or telomerase RNA, potentially useful	
CC	therapeutically, also to raise specific antibodies useful in	
CC	immunossays and therapeutically as inhibitors. Also contemplated	
CC	are transgenic animals in which the TP2 gene has been inactivated or	
CC	is overexpressed. TP2 polypeptides are administered i.v., s.c. or	
CC	orally, or they are delivered from engineered cells or gene therapy	
CC	vectors.	
SQ	Sequence 949 AA;	
SQ		

```

Query Match      47.3%; Score 2790.5; DB 1; Length 949;
Best Local Similarity 60.8%; Pred. No. 7.7e-250;
Matches 576; Conservative 106; Mismatches 217; Indels 49; Gaps

QY      1 MTRAPRCVAVSLLRSRYREWVPLATFVRRLGPEGRRLVQFGDKPIRYTVLAOCLVCMHW 60
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      23 MPAPRCRAVRSLLRSRYREVLPLATFVRRLGPGQWRLVQRGDPAAFRALVAOCLVCPW 82
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY      61 GSQPPADLDFHQVSSLKELVARVQRLCERNERNVLAFFGELLNARGGPPMAFTSSVR 120
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      83 DARPPAAPSFQVSCLELVARVQRLCERCAKNVLAFFGALLDARGGPPPEAFTTSVR 142
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY      121 SYLNTVETIYVSGAWMLLSRVGDDLVYLLAHCALVLLVPVSCAYQVCGSPLVQICA 180
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      143 SYLNTVTDALRGSGAWGLLLRRVGGDVLVHLARCALFVLVARSCAYQVCGPPLVQLGA 202
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY      181 TTDTPWSVASRYRTPRVGRNTNLFRLQIKSSSRQEPAPKPLALPSRGTKRHLSTSTS 240
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      203 ATQARPPPHAS-GPRRLG-----CERAWNHSVREAGVPLGLPALPGARRRGGSASRS 253
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY      241 VPSAKKARCYPVPRVEEGP-----HRQVLPTPSGKSW-VPSPARSPEVPTAEKDLSSK 292
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      254 LPLFKPRRGAAPPEPTVPVGGSWAHFGRTGSPDRGFCVVSFAR-----PAEATSLE 308
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY      293 GKVSDSLIS-GSVCCCHKKSSTSLSPRONAQFLRP-FIETRHFLYRGDGOERINPSF 350
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

Db 309 GALSGRTHSHSPVGRQHAGHPSTSRPPRWDTPCPPVYAETKHFLYSSGD-KEQLRPSF 367
 QY 351 LLSNQLPNLTGARRLVEIFLGSRPRTSGPLCRTHRLSRVYWMWRPLFOOLLVNAECQY 410
 Db 368 LLSLSRPSLTGARRLVEIFLGSRPWPMPGTPRRLPLRQIYWMWRPLFLELLGNAOCQY 427
 QY 411 VLLRSHCHFRFTANOQVTDAL-----NTSPPHLMDLLRLHSSPWPQY 452
 Db 428 GVLLTHCPFLRAA--VTPAAGVCAREKPGCSVAAPPEEDTDPRRLVQLLRQHSSPWPQY 484
 QY 453 GFLRACLCKVYSASLWGRTHNRERFEKNLAKETSLGKYCKLSLOELWKMVKVEDCHWLRS 512
 Db 485 GEVRACLRLVPPGLWGRSHNRERFLNTKKFISLGRHAKLSLOELTWKMSYRDCAWLRR 544
 QY 513 SPGRDRVPAAEHRRLRERILATFLFWMLDYTVVQLLSFFYITESTFQKNRLFYRKYSVMS 572
 Db 545 SPGVGCVPAAEHRRLREITLAKFLHLMVSVYVWELLRSFFYVETTFQKNRLFYRKYSVMS 604
 QY 573 KIQSTGVQHLRLVRLRELSEOEVRHHQDPWLAMPICRLRFTPKPNGLRPIVNMYSMGT 632
 Db 605 KQSIGIGIQHLKRVQLRELSEAEVRQREARPAALLTSRLRFTPKPDGURPIVNMDDYVYGA 664
 QY 633 RALGRKQAOHQHTORLKTFLSMLNIERTKHPHLMGSSVLGMNDIYRTWRAFLVLRALDQ 692
 Db 665 RTFREKRAERLTSRVKALFSVLNYERARRPGLLGASVLGLDDIHRANRTFVLVRAQDP 724
 QY 693 TPMAFYKADVGTGAYDAIPQGLKLVVVVANNRHSSSTYCIROYAVVRRSDGQVHKFSRR 752
 Db 725 PPELYFVKVDVGTGAYDIPQDRLTVEVIAIIRK-PQNTCYRRYAVVQAAHGHVRKAFKS 783
 QY 753 QVTLSDLPQMGQFLKHLQSDASALRNSVVEIQSIWNNESSLSLDFEFLHFLRHSVVK 812
 Db 784 HVSTLTDLPYMQEVAHLQET--SPLRAVVEIOSSUNEAASSLDFVFLRFMCHHAVR 841
 QY 813 IGDRCYTCOCGIPQGSSTLLCSLCFCGDMENKFLAEVORDGLLRFVDDFLVTPHLQD 872
 Db 842 IRGASYVQCCGIPQGSSTLLCSLCYGMENKFLAEVORDGLLRLVDDFLVTPHLTH 901
 QY 873 AKTEFLSLVHGVEYCGMNLQKTVNFPVEPTGLGGAAPQLPAHCL 920
 Db 902 AKTEFLTLVRGVEYCGVNLKRTVNFVEDEALGGTAFVQMPAHL 949

 RESULT 12
 W46997
 ID W46997 standard; Protein; 807 AA.
 AC W46997;
 DT 13-AUG-1998 (first entry)
 DE Human telomerase reverse transcriptase Delta182 variant.
 KW Human; telomerase reverse transcriptase; hTERT; TRT; diagnosis;
 OS prognosis; cell proliferation; cancer; ageing; ribonucleoprotein.
 OS Synthetic.
 OS Homo sapiens.
 PN G2317891-A.
 PD 08-APR-1998.
 PF 01-OCT-1997; 020890.
 PR 14-AUG-1997; US-315503.
 PR 01-OCT-1996; US-724643.
 PR 18-APR-1997; US-844419.
 PR 25-APR-1997; US-846017.
 PR 06-MAY-1997; US-851843.
 PR 09-MAY-1997; US-854050.
 PR 14-AUG-1997; US-911312.
 PR 14-AUG-1997; US-912351.
 PA (GERO-) GERON CORP.
 PA (OYTE-) UNIV TECHNOLOGY CORP.
 PI Andrews WH, Cech TR, Chapman KB, Harley C, Lingner J,
 PI Morin GB, Nakamura T, Harley CB;
 DR WPI; 98-171633/16.
 DR N-PSDB: V22382.
 PT Pure and recombinant human Telomerase Reverse Transcriptase and its
 PT variants - are useful in the diagnosis, prognosis and treatment of

cell proliferation conditions especially cancer and ageing
 Disclosure; Fig 20; 387pp; English.
 The present sequence represents a human telomerase reverse transcriptase (hTERT) variant from the present invention. The present invention also describes the following methods: (A) determining whether a test compound is a modulator of hTERT, by detecting the change in hTERT recombinant protein or polynucleotide, on administration of the compound; (B) preparation of recombinant telomerase by contacting a protein the hTERT RNA or protein in a sample by binding a relevant probe to the sample and detecting the complex formed or in the case of RNA detection, amplifying the product and correlating the presence of complex or amplification product with presence of hTERT in the sample; and (D) increasing the proliferation of a vertebrate cell by increasing hTERT expression; and (E) the use of an agent that causes an increase in cell vertebrate cell proliferation to create a medicament that inhibits ageing. A protein preparation of hTERT and the polynucleotide encoding hTERT can be used in the manufacture of medicaments for inhibiting the effect of ageing or cancer. Inhibitors of telomerase activity can be used to treat conditions that are associated with high telomerase activity. A protein preparation of hTERT can also be used in the new methods.
 Sequence 807 AA;

Query Match 37.2%; Score 2195; DB 1; Length 807;
 Best Local Similarity 58.3%; Pred. No. 1.1e-194;
 Matches 462; Conservative 88; Mismatches 191; Indels 52; Gaps 12;

QY 1 MTRAPRCVAVRLSRYSREVWPLATFVRLGPGRLVQPGDKIYRILVQAOLCVCMHW 60
 DB 1 MPRAPRCVAVRLSRYSREVWPLATFVRLGPGRLVQPGDKIYRILVQAOLCVCMHW 60
 QY 61 GSQPPADLSEHQVSKELVAVRVQRLCERNENYLAFCGELLNARCGPPMAFTSSVR 120
 DB 61 DARPPAAPSPVQVSKELVAVRVQRLCERNENYLAFCGELLNARCGPPMAFTSSVR 120
 QY 121 SYLNTVETLRVSGAWMLLSRVGDDLLVLAHACALYLLVPPSCAYQVCGSPLYOICA 180
 DB 121 SYLNTVETLRVSGAWMLLSRVGDDLLVLAHACALYLLVPPSCAYQVCGSPLYOICA 180
 QY 181 TTDIWPVSASVTRPVGRNFTNLFLOQIKSSROEAPKPLAPSRGTKRHLSTSTS 240
 DB 181 ATQARPPPHAS-GPRLRLG-----CERAWNHSVREAGVPLGAPGARRRGGASRS 231
 QY 241 VPSAKKARCYVPRVEEGP-----HROVLPSPGSKW-VPSPARSPVEPTAEKDLSSK 292
 DB 232 LPLKRRPRGAAPERTPVQCGSWAHGPGTRGSDRGFCVVSFAR-----PAEATSL 286
 QY 293 GKVDLSLS-GSVCKHKPSSTLSLSPRONAFOLRP-FIETRHLYSRGQGERLNPSF 350
 DB 287 GALSOTRSHSPVGRHQHAGPSTSRPPRWDTPCPVYAEKHFYSSGD-KEOLRPSF 345
 QY 351 LLSNLPNLTGARRLVEIFIGSRPTSGPLCRTHLSRRYWMRPLFOOLLVNAECQY 410
 DB 346 LLSLRPSLTGARRLVEIFIGSRPWPMPGTPRPLPQRYWQMRPLFLELLGNAOCY 405
 QY 411 VLLRSHCRFRANQQVDAL-----NTSPHMLDLKLSHPQVY 452
 DB 406 GYLLKTHCPLRRA---VTPAAGVCAKREKPGSVAAPAEEDTDPRRLVQLLRQHSPPQVY 462
 QY 453 GFLRACLVKVSASLWGRHNRFRFKNLKFISLGKYGKLSLOELMKMKVEDCHWLS 512
 DB 463 GFVRACLRPLVPGWLSGSHNRFRFRFKNLKFISLGKHAQLSLQELTWKMSVRDCAWLR 522
 QY 513 SPCKDVRPAEHRRLREILATFLWMDTYVYVQLLRSFYITESTFQKNRFFFYKKSWS 572
 DB 523 SPVCGVPAEHRRLREILATFLWMDTYVYVQLLRSFYITESTFQKNRFFFYKKSWS 582
 QY 573 KLOSGVGRHLRVRRELQSEVRRHQDTWLAMPICRLRFPKPNGLRPVNMYSMGT 632
 DB 583 KLOSGVGRHLRVRRELQSEVRRHQDTWLAMPICRLRFPKPNGLRPVNMYSMGT 642

QY 633 RALGRKQAOHFTQRLKTLFSLMNTERTKHPHMGSSVLGMNDIYRTWRAFLVRLALDO 692
 DB 643 RTFREKRAELRTRSVKALFSLVLYERARRPGLGASVLGLDDIHRWAFTRFVLRVAQDP 702
 QY 693 TPRMYFKADVTCAYDAIPQGLVEVAVNMIRHSESTYCIROYAVVRRDSQGVHKSFR 752
 DB 703 PPELYFKVDVTGAYDTIPQDRLEVIASLIK-PQNTYCVRYAVVQKAHGHVRKAFKS 761
 QY 753 QVTTSLDLPQYM 765
 DB 762 HV-----LRVPG 769
 RESULT 13
 W97384
 ID W97384 standard; Protein; 591 AA.
 AC W97384;
 DT 14-MAY-1999 (first entry)
 DE A catalytic telomerase protein.
 KW Catalytic telomerase; diagnosis; disease; telomerase activity.
 OS Homo sapiens.
 PN J11046768-A.
 PD 23-FEB-1999.
 PF 01-AUG-1997; 207708.
 PR 01-AUG-1997; JP-207708.
 PA (MITU) MITSUBISHI CHEM CORP.
 DR N-ESDB; X15923.
 DT New catalytic protein of telomerase of a higher animal and a gene coding it - useful for diagnosis of diseases caused by the change in activity of a telomerase.
 PS Claim 1; Page 11-14; 18pp; Japanese.
 CC The present sequence represents a catalytic telomerase protein.
 CC The products are useful in drug compositions for the diagnosis of diseases caused by the change in activity of telomerase.
 SQ Sequence 591 AA;

Query Match 35.1%; Score 2073; DB 1; Length 591;
 Best Local Similarity 67.7%; Pred. No. 1.3e-183;
 Matches 402; Conservative 72; Mismatches 114; Indels 6; Gaps 3;

QY 532 ATFLFWLMDTYVYVQLLRSFYITESTFQKNRFFYKRSVWSKLSIGVROHLRVLREL 591
 DB 1 AKFLHWSVYVYVQLLRSFYITESTFQKNRFFYKRSVWSKLSIGVROHLRVLREL 60
 QY 592 SOEEVRHODTWLAMPICRLRFPKPNGLRPVNMYSMGTALGRKQAOHFTQRLKTL 651
 DB 61 SEAEVRQREARPALTSRLRIPKDPGLRPVNMIDYVVGARTFREKRAELRTRSVKAL 120
 QY 652 FSLMNTERTKHPHMGSSVLGMNDIYRTWRAFLVRLALDOPTPRMYFKADVTCAYDAIP 711
 DB 121 FSVLYERARRPGLGASVLGLDDIHRWAFTRFVLRVAQDPPELYFKVDVTGAYDIP 180
 QY 712 QGKLVEVAVNMIRHSESTYCIROYAVVRRDSQGVHKSFRROVTTLSLDLPQYMGOFLKHL 771
 DB 181 QDRLEVIASLIK-PQNTYCVRYAVVQKAHGHVRKAFKSHVSTLTDLPQTRFVRL 239
 QY 772 QDSASALNSVYVQISISNNESSSLDFFLHRLHRSVVKIGDRCYTCOCGIPGSSLS 831
 DB 240 QET--SPLRDVAVYVQISISNNESSSLDFFLHRLHRSVVKIGDRCYTCOCGIPGSSLS 297
 QY 832 TLLCSLCFGDMENKLPFAVORDGLLRLRVDDFLLVTPHLDQAKTFLSTLVHGVPEYGCMI 891
 DB 298 TLLCSLCYGD MENKLPFAVORDGLLRLRVDDFLLVTPHLDQAKTFLSTLVHGVPEYGCMI 357
 QY 892 NLQKTVVNFVEPCTGGAPYOLPAHCLFPWCGLLDTPLEFVCDYSGYAGTSTIKSL 951
 DB 358 NLKRTVYVNFVEDEALGCTAFVQMPAHGLFPWCGLLDTPLEFVCDYSGYAGTSTIKSL 417
 QY 952 TFQSVFVAGKAGTMRNKLRLVLRKCHGLFLDLQVNSLQTCVINYKIFLQAYRFAHVQ 1011
 DB 418 TFNGRAGRMRNKLRLVLRKCHGLFLDLQVNSLQTCVINYKIFLQAYRFAHVQ 477

Pure and recombinant human Telomerase Reverse Transcriptase and its

QY 539 MDTYVVQLLRSEFFYITESTFOKNRIEEYPKSWSKI OCTUBROU

Mon Aug 16 10:21:26 1999

PT variants - are useful in the diagnosis, prognosis and treatment of
PT cell proliferation conditions especially cancer and ageing

PS Example 6; Page 224; 387pp; English.

CC The present sequence represents a fusion protein from an example

CC of the present invention which describes human telomerase reverse

CC transcriptase (hTERT). The present invention also describes the

CC following methods: (A) determining whether a test compound is

CC a modulator of hTERT, by detecting the change in hTERT recombinant

CC protein or polynucleotide, on administration of the compound;

CC (B) preparation of recombinant telomerase by contacting a protein

CC preparation of hTERT with a telomerase RNA component; (C) detection

CC of the hTERT RNA or protein in a sample by binding a relevant

CC probe to the sample and detecting the complex formed or in the case of

CC RNA detection, amplifying the product and correlating the presence of

CC complex or amplification product with presence of hTERT in the sample;

CC and (D) increasing the proliferation of a vertebrate cell by increasing

CC hTERT expression; and (E) the use of an agent that causes an increase in

CC cell vertebrate cell proliferation to create a medicament that inhibits

CC ageing. A protein preparation of hTERT and the polynucleotide encoding

CC hTERT can be used in the manufacture of medicaments for inhibiting the

CC effect of ageing or cancer. Inhibitors of telomerase activity can be

CC used to treat conditions that are associated with high telomerase

CC activity. A protein preparation of hTERT can also be used in the new

CC methods.

CC Sequence 538 AA;

SQ

Query Match 18.5%; Score 1092.5; DB 1; Length 538;
Best Local Similarity 71.2%; Pred. No. 1.1e-92;
Matches 215; Conservative 26; Mismatches 58; Indels 3; Gaps 1;

QY 824 IPQSSILTLCLSCFGDMENKLFQAEVQDGLLRLRFVDDLLVTPHLDQAKTFLSTLVHG 883
DB 237 IPQSSILTLCLSCYCGDMENKLFQAEVQDGLLRLRFVDDLLVTPHLDQAKTFLSTLVHG 883

QY 884 VPEYGCMLNLTQVYVFPFVETGLGGAAPYOLPAHCLFPWCGLLDQTOLEVCDSYGA 943
DB 297 VPEYGCMLNLTQVYVFPFVETGLGGAAPYOLPAHCLFPWCGLLDQTOLEVCDSYGA 943

QY 944 QTSIKTLTFOSEVFKAGTKMKNLLSVLRKCHGLFLDLQVNSLQTVCTNIYKIFLIQAY 1003
DB 357 RFSIRASVTFNRFAGNMRKRLFGVLRKCHSLFLDLQVNSLQTVCTNIYKIFLIQAY 416

QY 1004 RFHACVQLPFDQVRKKNLTFFLGISSQASCCYAILKVNPGMTLKASG---FPPEAA 1060
DB 417 RFHACVQLPFEQVQWKNPTFFLRVISTASLCYSILKAKNAGSLGAKGAGLPSEAV 476

QY 1061 HMLCYQAFLLKLAHSVYKCLGLPLRTAOKLCKRLPEATMTILKAAADPALSTDFQTI 1120
DB 477 QWLCHOAFLLKLRHRTVYVPLGLSLRTAQQLSKRLPGTTLTAAEAANPALPDEFKI 536

QY 1121 LD 1122
DB 537 LD 538

Search completed: August 13, 1999, 21:34:53
Job time: 999 sec

Result No.	Score	Query Match %	Length	DB	ID	Description
1	341	5.8	884	2	S53396	telomerase catalytic
2	131	2.2	778	2	I38487	tastin - human
3	125	2.1	1523	2	T02730	probable reverse t
4	121	2.1	2088	2	EY1436	hypothetical prote
5	111.5	2.0	21461	2	E71461	proline-rich prote
6	111.5	1.9	1415	1	EDBEGA	immediate-early pr
7	111.5	1.9	2606	2	T03159	large legume
8	110	1.9	907	2	A24938	hypothetical T2 pr
9	109.5	1.9	1214	2	T00356	hypothetical prote
10	108.5	1.8	351	2	S07554	hypothetical prote
11	108	1.8	880	2	S49627	regulatory protein
12	107	1.8	1456	2	T01397	LTR transposon pol
13	106.5	1.8	1256	2	C71436	probable resistance
14	106	1.8	1115	2	A45761	Ca2+-transporting
15	106	1.8	584	2	S40013	hypothetical prote
16	105.5	1.8	3898	2	S57437	genome polyprotein
17	104.5	1.8	916	2	A38418	jockey protein 2 -
18	103.5	1.8	536	2	E70066	hypothetical prote
19	103	1.7	508	2	S74977	gag-akt polyprotei
20	102.5	1.7	763	1	A40831	probable zinc-fing
21	102.5	1.7	1094	2	T00814	probable membrane
22	102.5	1.7	1556	2	S59393	probable permeal
23	102.5	1.7	781	2	A37956	sulfate permease I
24	102	1.7	808	2	S07649	gene coI intron 1
25	102	1.7	850	2	S56015	gastric mucin - hu
26	101.5	1.7	884	2	T01041	hypothetical prote
27	101	1.7	994	2	I49376	c-met tyrosine kin
28	101	1.7	1004	2	T00795	tRNA-processing kn
29	100.5	1.7	1684	2	T02632	hypothetical prote
30	99.5	1.7	360	2	S18458	tubulin beta chain
31	99.5	1.7	810	2	S49744	AMP deaminase (EC
32	99.5	1.7	968	2	T00322	hypothetical prote
33	99.5	1.7	528	2	I47141	gastric mucin (Clat
34	99	1.7	1680	2	T01367	hypothetical prote
35	99	1.7	1228	2	S59681	probable membrane
36	99	1.7	840	2	S69204	pheromone response
37	99	1.7	1621	2	S63526	MAP-185 protein -
38	99	1.7	646	2	JC5583	85k SH3 domain-con
39	98.5	1.7	3898	2	S59295	polyprotein - hog

polyprotein - hog

A;Cross-references: SGD:S0004310; MIPS:YLR318w
A:Map position: 12p

Query Match	5.8%;	Score 341;	DB 2;	Length 884;
Best Local Similarity	19.8%;	Pred. No. 6e-18;		
Matches 208;	Conservative 148;	Mismatches 355;	Indels 338;	Gaps
QY	33	PEGRRLVOPGDKPIYRTLVVAOCLVCMHWGSO	PPADISFHOVSSLKELVARVVQRLCERN	92
Db	45	PNSRKIALP-----CL-----	PGDLSHRKV-----IDHCIIIVLLTGEL	77
QY	93	ERNVLAFGELL-NEARGPPMAFTSSVRSYLPNTVIETLRVSCAWMLLSRVGDDLLVY	151	
Db	78	YNNVLTEGKYIARNEDVYNSLFCISAVNV-----	VTLKGAAKMFHSLVGTVAEVD	129
QY	152	LLAHACALLVPPSCAVQVCSPLYQI-----CATTDIWP-----	SVSASYRTPRPVGR	200
Db	130	LLIN---YTVI-----QFNGQFTQIVGNRCNEP	LPKWKVORSSSSATAQIKQLTE	180
QY	201	NFTNRLRFQIQIKSSRQEAEPKPLAPSRGTRKHLSTLSTSVPSAKKARCIYPPRVSEGPH	260	
Db	181	PVTNKQFLHK-----	LINSSSF-----FPY	201
QY	261	ROVLPTPSGKSWPSPARSPVEVPTAEKDLSSKGVSDLSLGSVCCKHKPSTSLSPRP	320	
Db	202	SKILPSS-----SSIKKLTDL-----	REAIFTNLVKIP-	231
QY	321	QNAFQLRPFITRHFVLYSGDGOERLNPFLSLNQPNLTGARRLVEIIFLGSRPRTSGP	380	
Db	231	-----QRL-----	KVRNLT-----	241
QY	381	LCRTHLSRRYWMRPLFOOLLVNHAECQYVRLLRASHCRFTANQVTDALNTSP	---H	437
Db	241	-----LQKLLKRHLNLYSVILNSIC-----	PPLEGT	267
QY	438	LMDLRL--HSSPMQVYGLRACLKVVYSASLWGRTHNERFFKNLKFVILSGYKGLSL	495	
Db	268	VLDLSHLRQSPKERVKFIIVILQKLPQEMFGSKNGKIKIKNLLSLLPLNGVLPF	327	
QY	496	QELMWKMKVEDCHWLRSRSPGKDRVPAAEHLRL--ERILATFLWMDTYVVOLLRSFYI	553	
Db	328	DSLKKLRLLKDFRWLFISD-----INFTHNENLNOLAICFISFLQRLPIKPIQTFFYC	383	

Mon Aug 16 10:21:27 1999

554 TE-STFQKNLFFYKRSVWSKLOSIGVROHLERVLRELQBEV--RHHDWTWAMPIC 609
 384 TEISSTV---TIVFRDHTWKNLTPFIVEFKY-----LVENVCHNHSYTLNFNHS 436
 610 RLRETPKP--NGLR-----PIVNMYSMGTRALGRKKAQHTORLKTFLFSLNVERTKHP 663
 437 KMRIPKSNNEFRILAIAPCGADEEETIYKENHNAIQTKI-----LEYLRNKR 490
 664 HLMGSSVLGMNDIYTWRAFVLRV--RALDQTPRMFYFKADVTGAYDAIPQKLVVVAN 721
 491 TSF-TKIYSPTQIADRIKEFKORLLKFNKFNVLPELYFMKFDVKSCYDIPRMECMRLKD 549
 722 MIRHSESTVICROY-----AVVRDSOGQVHKFSFRQVTTLSLDLOPYMGFLK 769
 550 ALKNENGFFVRQYFNTWTGVLKFNVNASRPVPELYIDNVRTV----- 598
 770 HQDSDASALRNSVIEQISMNESSSSLFDFLHFLRHSVVKIGDCRYTCOGIPOGSS 829
 598 HLSNQDV-----INVMEIF-----KTALWVEDKCIYREDGLFQSS 635
 830 LSTLLCSLCEGDM---ENKLEAEVORDGLLRFVDDFLVTPHLDQAKTFSTLVHGVPE 886
 636 LSAPVDLVDDLEFYSEKASPSQDTLLKLADDFLIISTDQOQVINIKKLAGMGFQK 695
 887 YGCMINLQKTVVNPVPEPTLGAAPYQLPAHCLFPWCGLLLDTQTLEVFCDYSGYAQTS 946
 696 YNAKANRDKIL-----AVSQSDDDTIVIOFCAMHIFVKEVWKHSS----- 738
 947 IKTSLTQSFVKAGTKMRNKLVLKCHGLFDLQVNSLQVFCINI-----YK 996
 738 --TMNNEHRSKSGKIFRSIALFNTRISYKIDTNLSTNTVLMQIDHVVKNISECYK 795
 997 -----IFLQAYRHFACVIQPFDDRV 1018
 796 SAFKDLISINTQNMQFHS-----FLQRI 818

RESULT 2
 138487
 tasin - human
 C:Species: Homo sapiens (man)
 C:Date: 16-Feb-1996 #sequence_revision 16-Feb-1996 #text_change 16-Feb-1996
 C:Accession: 138487
 R:Fukuda, M.N.; Sato, T.; Nakayama, J.; Klier, G.; Mikami, M.; Aoki, D.; Nozawa, S.
 Genes Dev. 9, 1199-1210, 1995
 A:Title: Trophinin and tasin, a novel cell adhesion molecule complex with potential inv
 A:Reference number: 138487; MUID:95278733
 A:Accession: 138487
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-778 <RES>
 A:Cross-references: EMBL:U04810; NID:905355; PID:905356

Query Match 2.2%; Score 131; DB 2; Length 778;
 Best Local Similarity 22.4%; Pred. No. 0.061;
 Matches 107; Conservative 44; Mismatches 167; Indels 160; Gaps 26;
 2 TRAPCPAVRLSLRSRYREVWPLATEVRRLGPGRRLVOPGPKYIRLVA-----QC 54
 358 TNPSTPRVQQ-----AQLRGVSPQS-----CSEDPAFPWEQVAVRLFDQESC 401
 55 LVCMHNGSQPP---PADLSFHQVSSKLKELVAR---VVORLCERNERNVLAFAFELLNEAR 108
 402 IRSLESGKPPVATPSPHNSRNPSPLOEVKIQIGLQQLLQKQEVGLVGGCVPLN--- 459
 109 GGP--PNAFTSSVRSYLPNTVIETLRVSGAMLLLSRVGDDLLVYLAHICALYLLVPPSC 166
 459 GGSLLDWELOPLTEISRLNATEHNSGT-----SHLPGLLKHSGL-----PKPC 504
 167 AYQVCGSLPVOICATTDIWP-----SVSASVYRTPRPVGRNFTN-L 205
 505 LPEECGEP--QPCPPAEPGPPAFCRSEPEIPESLQEQLEVPPEYPPAEP--RPLESCC 560

206 RFLQIKSSRQE-----APKPL-----ALPSRGTKRHLSTSTSVPSAKK 246
 561 RSEPEIPSSQEQLEVPPECPPAEPRLSESYCIEPEIPSSRQEOL-----EVPE--- 613
 247 ARCYVPVVEEGPHROVLPTPSGKSWPSPARSPEVTAEKDLSSKGVSDLSLGSV-- 305
 613 ----PCPPAEPGP---LQPSQGGSGPPGC-----PRVELGASEPCTLEHRSLESSLPP 660
 305 CCKH-KPSTSL-----SPRONAFOLRFIETRHFLYSRGQGERLNPFSLLSN 354
 661 CCSQWAPATTSILFSSQHPICASPICSLQSLRP-----PAGQAG-----LSN 703
 355 LOPNLGTGARLV-----EILFGRSPRTSGP--LCFTHLSRRYQ 393
 704 LAPRTLALRESLKSCLTAHCFHEARLDDECAFTSRASPSGPTRVCTNPVATLEWQ 761

RESULT 3
 T02730
 probable reverse transcriptase T914.6 - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 24-Mar-1999
 C:Accession: T02730
 R:Rounsley, S.D.; Lin, X.; Kaul, S.; Shea, T.P.; Fujii, C.Y.; Mason, T.M.; Shen, M.;
 submitted to the EMBL Data Library, August 1998
 A:Description: Arabidopsis thaliana chromosome II BAC T914 genomic sequence.
 A:Reference number: Z14710
 A:Accession: T02730
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-1529 <ROU>
 A:Cross-references: EMBL:AC005315; NID:g3461834; PID:g3461840
 C:Genetics:
 A:Map position: II
 A:Introns: 428/1; 532/3
 A:Note: T914.6

Query Match 2.1%; Score 125; DB 2; Length 1529;
 Best Local Similarity 19.2%; Pred. No. 0.48;
 Matches 209; Conservative 147; Mismatches 341; Indels 394; Gaps 58;
 142 SRVGDDLLVYLAHICALYLLVPPSCAYQVCGSPYQICATDIWPSVSAKYRTPRVGRN 201
 220 SETGYDAMVEY-----KYPWLPSPRSCSKWGHIOEVLCTR---PSPQLSTPT--- 266
 202 FTNLRFLQIKSSSRQEPAP---KPLALPSRGTKRHLSTSTSVPSAKKARCYVP--PRV 255
 266 -----EIEETEDKTEPPLMKERPLEILSKSPSATLTKTLNGDSHTQKV---PMKNPTV 314
 256 -----BEGHROVLPTPSGKSWPSPAR-----SP-----EYPTA 285
 315 LQNKGEVAEENEENLKDEGPHWTVSPAKVGRQSPQOEYVNNHASPSPRFDVLAVEDTA 374
 286 EKDLSSKGVSDLSLGSVCCCKHKPSTSLSPPRQNAFOLRPFET---ETRHFLYSRGD 341
 375 DTNGHEGEIVVLQ-----QLNDNSDV-----GVQGTPEYIPRISKSHQKVKVSSA 422
 342 GQE-----RLNPSFLLSNL-----QPNLTGARRLVEI 368
 423 NQNTKGIWVWRDNARLTPEVKSSQMITCSILLGKEEEFFCFSFIYASNFVEER---I 479
 369 IFLGSRPTISGGLCRTHRLSRRYWOMRPLFOQLLVNHAECQYVRLLRSHCRFRATANQVT 428
 480 LWEDIRSHHDSPIR-----NKPWILCGDFNEILGEGHSNY----- 517
 429 DALNTSPHMLDLRLHSSPWQYVFLRACLCKVVSASLWCTRHNERREFKNLKKFISLG 488
 517 DNSPYTPPGMRD-----FQBIGRL-----MLEAATGGR---KPF-----KFNWV- 554
 489 KYGLSLQELMKWKVEDCHWLSSP-----GKDRVPAAE 523

Db 554 -----LTKLPQPLVSVSHWASAPLYVSTALYRESKLLKTLKPHILRELKGLDLP 607
QY 524 HRLRE-RIL-----ATFLWMDTYVQLLRSFFYIT-----ESTFOK----- 561
Db 608 KRTREAHILCEKQATTLANPSOETIAEELKAVTDMTHLSELEEGFLKQSKLHWNVGD 667
QY 561 -NRLFFYRKSVKQSIGVQRLHVRRLREL-----SDEEVRHHODTWLAMPICR 610
Db 668 GNSYFHKAQVRKM-----RNSIREIRGPNATLQTSSEIKGEAREFN----- 713
QY 611 LRIPIKPN-----LRPIVMSYMGTRALGRKQQAQHTQRLKTLFSLMNYERTK 661
Db 713 -EFLNQSGDFHGISVEDLANLSYRCSVTDQNLITREVTGEEIQ--KVLFAFPN---NK 766
QY 662 HPHLMGSSVGLMNDIYR-TWRA-----FVLVRALDQTPRMFYKADVTGAYDAIPOGKLV 716
Db 767 SPQPDGYT-----SEFFKATWSLTPGDFIAAIQS-----FFVKG-----FLPKGLNA 808
QY 717 EVVANNIRHSES-----TYCIRQYAVVRDSDQGVHKSFRROVTTLSLDLPYMGOF 767
Db 809 TILALIPKDEATEMDYRPISCNVLYKISKILAN-----RLKLLPSFILQNSAF 862
QY 768 LKHLQSDASALRNSV-----IQSISMNESSSLFD-----PFLHFL----- 807
Db 863 VBERLLMENVLLATELVKDYHKESVTPRCAMKIDISKAFDSVQWQFLLTLEALNPPET 922
QY 807 RH-----SVYKIGDRC--YTQCOGIPQSSLSLTLSCFCGDMENKL-FAEVR 852
Db 923 RHWIKICISTAFSVQVNGELAGFFGSSRGLRGCCALSPYLFCVCMNVLSHMDIAAHR 982
QY 853 D-----GLL-LRFVDDFLL-VTPHLDQAKTFLSLTVHGVPEYGCWMLNQ----- 896
Db 983 NIGYHPKCEKIGLTHCFADDLVDFVGHQWSTEGVINVEKFAFGSGLOISLEKSTIYL 1042
QY 896 -----TVNRPVECTLGGAAPYOLPAHCLFPWCGLLDQTLTLEFVCDYSGYAO 945
Db 1043 AGVASDRVOTUSSPFFANG-----QLPVRYL-----GLPLTKQMT-ADYSPLE- 1089
QY 946 SIKT---SLTFQSVKAGTMRKLLSVLRKCHGLFLDLQVNSLQTCVINYKIFLQOA 1002
Db 1089 AVTKISSWTARSLSVAGR-----LALLNSVIVSIAN-FWMSA 1125
QY 1003 YRFHA-CVIQL 1012
Db 1126 YRLPAGCIREI 1136
RESULT 4
E71436
Hypothetical protein - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
A:Variety: Columbia
C:Date: 03-Aug-1998 #sequence_revision 03-Aug-1998 #text_change 05-Dec-1998
C:Accession: E71436
P.: Bevan, M.; Bancroft, I.; Bent, E.; Love, K.; Goodman, H.; Dean, C.; Bergkamp, R.; Dirk
P.: Wedler, H.; Wedler, E.; Wambutt, R.; Weitzenecker, T.; Pohl, T.M.; Terry, N.; Giel
avanagh, T.; Hempel, S.; Kotter, P.; Entian, K.D.; Rieger, M.; Schaeffer, M.; Funk, B.
Nature 391, 485-488, 1998
A:Authors: Mueller-Auer, S.; Silvey, M.; James, R.; Montfort, A.; Pons, A.; Puigdomenech
erhoft, A.; Moores, T.; Jones, J.D.G.; Eneva, T.; Palme, K.; Benes, V.; Rechman, S.; Ans
C.: Chalwatzis, N.
A:Title: Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of Arabidopsis thal
A:Reference number: A71400; MUID:98121113
A:Accession: E71436
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-2088 <BEV>
A:Cross-references: GB:297342; NID:g2245031; PID:e327020; PID:g2245044
C:Genetics:
A:Map position: 4COP9-4G3845

Query Match

2.18; Score 121; DB 2; Length 2088;

Best Local Similarity 19.9%; Pred. No. 1.5;
Matches 151; Conservative 82; Mismatches 251; Indels 276; Gaps 37;
QY 88 LCERNERNVLAFCFFELLNEA---RGPPNAFTSSVRSYLPNTV----- 128
Db 1293 LSERKRRHIVETGLTLTQASVPREYWPYFAAAV--VLINRMPTVLSMESPFQKLFGS 1350
QY 128 ---IETLRVSGA---WM-----LLSRVGDDLLVLLAHALY----- 160
Db 1351 KPNYERLVRFGCLCPWLPYPYTHNKLERSRCVFLGSLTQTAYLCEDFVEHKRLYTSRH 1410
QY 160 -----LLVP-----PSCAYQVCGSPY-QICATTDIWPVSASVYRTPRPVGRNFT 203
Db 1411 VVFEASFPFSLNTSQSLPTVTFEQSSPLVPLTSSSSVLPSCLS--PCTVLHQOQP 1468
QY 204 NURELQIKSSSRQAPKPLALPSRGTXRHLSTSTSVPS-----A 244
Db 1469 PVTTPNSPHSQPTTSPAPLS-PHRSTTMDQVQPTAPENGEPEPEAQSPPIGSLNPT 1527
QY 245 KKARCYPPRVEEGPHROVLPPTSGKSWVPSPARSPEVPTAEKDLSSKGKYSVLSGVS 304
Db 1528 HEAFIGPLPNRNPNTNIEPT-----PAPHPKPVKPTTTTTPNRTTSDAS----- 1576
QY 305 CCKHPSSTLSLPPRONAFOLRPFIEIRHFLYGRGQO-ERLNPFSLLSNLQNLGTAR 363
Db 1576 ---HQPTA-----PQON-----QHNKTRAKNNIKKPNTKFSLTATLPNRS--- 1614
QY 364 RLVEIFLGRSPRSGPLCRTHLSRRYWO-----MRPLFQQLLVNHA 406
Db 1614 -----PSEPTNVQALKDKKRFAMSDFEAQOQRNHTWDLVPHESQLLVG-- 1659
QY 407 ECQYVRL-----RSHCRFTANQO--VTDALNTSPPHLMDLLRL-----HSSP 448
Db 1659 -CKWVEKLYLPNGAIDKYKARLVAKGNOQYGVDAETFSVPIKSTIRLVDAVAKD 1717
QY 449 WQVYGLRACLCVVSASLWGTNRINERFFKNLKKFISLGKYKGLSLQELMKMKVEDCH 508
Db 1718 WEIKOL-----DYNNAFLQGLTTEE-----V 1738
QY 509 WLRSPG---KDRVPAAEHLRLERILATFLWMDTYVQLLRSFFYI-TESTFQKNR 564
Db 1739 YMAQPPGFIDKDR-PTHVCLRKAIYG--LQAPRAYMELKQHLNIGVFNVSLSASLF 1795
QY 565 FYKSVMSKIQSIGVROHLERVLRLSQRVRRHODTWLAMPICRLRFPKPNGLRPV 624
Db 1796 IY----WSDKSSID-----AVLTSIAE-----RF-----SIKDPT 1821
QY 625 NMSYMGTRALGRKQQAQHTOR--LKTLSMLNYERTK-----HPHLMGSSVLCMN 674
Db 1822 DLHYFLGIEAT-RTKQGLHLMQRYIKDLAKHNMADAKPVLTPLTSPKLTLLHGGTKLN 1880
QY 675 DI--YRTWRAFLVRALDQTPRMFYKADVTGAYDAIPQ 712
Db 1881 DASEYRS-----VGSIQIYLAFTRPDIYAVNRLSQ 1911
RESULT 5
S71461
proline-rich protein Rad9 - inky cap (Coprinus cinereus)
C:Species: Coprinus cinereus
C:Date: 09-Dec-1997 #sequence_revision 09-Dec-1997 #text_change 26-Feb-1998
C:Accession: S71461; S78194
R:Seitz, L.C.; Tang, K.; Cummings, W.J.; Zolan, M.E.
Genetics 142, 1105-1117, 1996
A:Title: The rad9 gene of Coprinus cinereus encodes a proline-rich protein required f
A:Reference number: S71461; MUID:96271528
A:Accession: S71461
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-2157 <SEI>
A:Cross-references: EMBL:U34998; NID:g1353389; PID:g1353390
A:Accession: S78194
A:Molecule type: mRNA

A:Residues: 1-2157 <SET>
C:Gene: radg
A:Introns: 65/1; 151/2; 433/3; 472/1; 685/3; 844/3; 878/2; 1080/1; 1114/3; 1182/2; 1246/2; 2054/3
C:Keywords: DNA repair; meiosis
F:607-616/Region: nuclear location signal

Query Match 2.0%; Score 115.5; DB 2; Length 2157;
Best Local Similarity 18.3%; Pred. No. 4.3;
Matches 213; Conservative 155; Mismatches 357; Indels 439; Gaps 54;

QY 70 SFHQVSS-LKELV-----ARVQRLCERNRVLAF-GPELLNEARGGPPMAFTSSVRS 121
Db 884 SIRTVSALLMQLVQTSNDVIRINARLEKERONALALKROESISDLNGQP-----KSDPE 938
QY 122 YLPNTVIETLRVSGAMW-----LLLSRVG-----DDL 148
Db 939 FLDNIDMEERLYGGGLESATKAATTIIFLNSRAGRKTTKNSNEAYRAIFDNLDL 998
QY 149 LVYLHACALYLLVPPSCAVCGSPLYQICATTDIWPVSASVYRTPRVGRNFTNL--- 206
Db 999 LVVL-----YWPENPAASLLLNIAISM 1019
QY 206 ---RFLQOIKSSR--QAPKPLALPSRGTKRHLSTSTSVPSAKKARCVYVPRVEEGPH 260
Db 1020 FVWSSLDVVKSNQADITNAKSMALD-----HLGVIAAIRSS-----ILKVQKDEDGTS 1069
QY 261 RQVLPTSGKSWPSP-----ARSPEV-----PTAEKDLSSKGKYSDLSSGVC 305
Db 1070 YRGLKPLDEVDRQSPVEAIRQVHGRSPDVAAHLCRSTEDQAVESARELTAAILGHEL 1129
QY 306 CKHKPSSTILLSPPRONAFOLRPFETRHFLYSRGDQERLNPFSLLSNLQPNLTGARR- 365
Db 1130 ASLRVNWLDHPQDEDLDR-----DSSKAL--SFGL-KLKTLALRGWKD 1173
QY 365 -LVEIIFLGSRPRTSGPLCRTHRLSRVYQWQMRPLFQOLLVNHAEQCVVLLRSHCRPTA 423
Db 1174 PATDVFVIGSQEEVS-----RIDRLSE-----EIGTIQSLR----- 1205
QY 424 NQVTDALNTSPHMLDLRLHSSPQWYVGLRA-----CLCKVYS--ASLWGTGRHNERRF 477
Db 1205 -----NSFQPLNLVLSALDAP---VIFWTKALRQALCOIYVTSATILGTA-SVRQG 1252
QY 478 FKNLKKFTSLGKYGKLSQELMWMKMKVEDCHWLRSPP-----GKORVPAE----- 524
Db 1253 IEN-----HLLDSSPAVRDAAVELICKYKYMIDSPVAGN 1285
QY 524 --HRLRERILATFLWMDTYVQVLLRSFFYITESTQKNLFFYKRSVWSKLOSIGVRQ 581
Db 1286 YVQKTAERMADTGL--AVRKRVIKLLKSYGVVDDAQRKIDI----- 1326
QY 582 HLERYRLRELQOE-----VRHODTWLAMPICRLRPIKPNGLRP----- 623
Db 1326 --SARLVLRNVEDDGVKDLAKTLEELWFPp-----LPPSGMKVYKPTSSSNPNQDKA 1377
QY 623 ---IVNMSYSMTGRALGRKQA-----OHTQORLTKL----- 652
Db 1378 ALLSKVAIIMGTAANFRDQRSPLEMDLHKIISDKEGDEAASLHORVABICETLIDGLVDA 1437
QY 652 -----FSMLNERTKH-----PHLMGSS-----VLGMNDIVRTWRAF 683
Db 1438 TDLPGFTIINCIRTLHLTAAYPSILPGTHASTLLPYLKNASTTEELTSDFLKFIRA- 1497
QY 684 VLVRALDQTPRMFEVK-ADVTGAYDAIPQGLVEVAMNIRHSESTYCIROYAVVRDSD 742
Db 1497 --SIPHMPKTPQIRAGIADVAALILKPFQGV-----NILQEAAGCMC-----AVVR--- 1543
QY 743 QGVHKSFRQVTTLSDOPLYMGQFLKHLQSDASALRNSVIEQISNMNESSSLDFDF 802
Db 1543 --HLTHDFKRLINLLKGCNARLLSYLRH-----PPTQQLNNV-----ESKTLMLLF 1587


```

QY 765 GQFLKHLQSDASALRNWVIEQSSSSLMNNESSSSLDFFLHLFRLHSVVKIGDRCYTQCGI 824
Db 1477 ----KLINDSEVSS-----KLTYSSEAA-----LHWESTCAHIA-----TQCYP 1514
QY 825 PQGSSLSLTLCSLFCGDMENKLFREYVORDGLLLRFVDLFLVTPHLDQAKTFLSTLVHGV 884
Db 1515 AVGAKLQAIIA-----LVAKVK-----PKLESLSV-LENOANT 1547
QY 885 PEYGMINLQKTVVNFVPEPTL--GGAAPYQLPAHCLFPMWCGLLLDLTQTLEVEDSYGA 943
Db 1548 DD-----INILQKALS-SLDPKRITGTSKVQ-----EWTQVKOLEKL-----LA 1587
QY 944 QTSIKTSL--TFQSVFKAGTMMN--KLLSVLRKKGHLF-----980
Db 1588 DTEASVWNIQLRHMLTARSTLLNLANLKOKTTSLEYKWKVKEHKTGPTSPITGTIKE 1647
QY 980 LDQVNSLOTVCINIKYIFLQAYRFH-ACVIQLPQDQVRKRNLFGLIISQASCCYA 1038
Db 1648 LDLYLT-----FKLKFIYYEYNQACIFS-----TFAPASKIDAKDAL 1687
QY 1039 ILKVKNPQMTLKASGFPP 1057
Db 1688 PLTPQSPVFEDNTSGSPSP 1706

RESULT 8
A24938
hypothetical T2 protein - human herpesvirus 4
C:Species: human herpesvirus 4, Epstein-Barr virus
C>Date: 02-Jun-1988 #sequence_revision 02-Jun-1988 #text_change 29-Jan-1999
C:Accession: A24938; S33010; I49034
R:Rodescott, M.; Brison, O.; Perricaudet, M.
Nucleic Acids Res. 14, 2611-2620, 1986
A:Title: An Epstein-Barr virus transcription unit is at least 84 kilobases long
A:Reference number: A24938; MUID:86176753
A:Accession: A24938
A:Molecule type: mRNA
A:Residues: 1-907 <BOO>
A:Cross-references: GB:X04060; NID:g59579; PID:g59580
A:Experimental source: strain B95-8
R:Farrell, P.J.
submitted to the EMBL Data Library, March 1988
A:Reference number: S32973
A:Accession: S33010
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-113 <FAR>
R:Cross-references: EMBL:V01555
R:Apolloni, A.; Moss, D.; Stumm, R.; Burrows, S.; Suhrbier, A.; Misko, I.; Sci
Eur. J. Immunol. 22, 183-189, 1992
A:Title: Sequence variation of cytotoxic T cell epitopes in different isolates
A:Reference number: A49034; MUID:9211623
A:Contents: nuclear antigen EBNA-3
A:Accession: I49034
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 278-302 <APO>
A:Cross-references: GB:S79227; NID:g242887; PID:g242888
A:Experimental source: human A-type strain B95-8
A:Note: sequence extracted from NCBI backbone (NCBIN:79227, NCBI:P:79239)
C:Superfamily: human herpesvirus 4 nuclear antigen EBNA-3A

Query Match
Best Local Similarity 1.9%; Score 110; DB 2; Length 907;
Matches 82; Conservative 48; Mismatches 138; Indels 100; Gaps 20

QY 38 LVQPGDPKPIYRTLVIAQCLVCMHWGQPPADLSFHSQVSSKLVELYRVVQRLCERNVL 97
Db 332 IVSRGGPKYKR-----PPI-----FIRLRLMLMRAGKTEQ--- 365
QY 98 AFGFELLNEARG---GPPMAFTSSVRSYLPNTVETIRVSGAMLLSLRVGDDLLVYLLA 154

```

D8 RFGSNEEALQCVLVL-IYPGLVGATVWS-----CSILALTHLLFEFKGL 851

QY 398 F-----QQLVNNAECQYVRLLRSHCRFTANQOYTDALTNTSPPHMLDLLRHSSPWQVY 452
 :||||| ||| | | | :
D8 852 MGTSTVEQLLEN-----VCILLASRTR-----DVVK-----SAL 880

QY 453 GFLRACL-----CKVYSASLWGTNRHERRFF-----KNL-KKFTSLGKYGKLSIQ 496
 ||| : ||| | | | :
D8 881 GFIKAVTVMDVAHLAKHVOLVWEATGKLSDMMRRHRMCLRNLTKFIR--KFG----F 934
 ||: : ||: : ||: :

QY 497 ELMWMKMVEDCHMLRSRGKDRVPAAEHR 525
 ||: : ||: : ||: ||
D8 935 ELVKRLPEEVHRVLRVNIRKAERAKRHR 963
 ||: : ||: : ||: ||

RESULT 10

S50754
hypothetical protein WP6 - Chlamydomonas eugametos
C:Species: Chlamydomonas eugametos
C>Date: 14-Jul-1995 #sequence_revision 21-Jul-1995 #text_change 09-Sep-1997
C:Accession: S50754
R:Woessner, J.P.; Molendijk, A.J.; van Egmond, P.; Kils, F.M.; Goodenough, U.W.; Hari
Plant Mol. Biol. 26, 947-960, 1994
A>Title: Domain conservation in several volvocalean cell wall proteins.
A:Reference number: S50754; MUID:95093034
A:Accession: S50754
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-351 <WOF>
A:Cross-references: EMBL:L29028; NID:g530877; PID:g530878

Query Match 1.8%; Score 108.5; DB 2; Length 351;
Best Local Similarity 24.0%; Pred. No. 0.97; 144; Indels 85; Gaps 13;
Matches 81; Conservative 28; Mismatches 28

QY 104 LNFARGGPMAVTSSVRSYLPNTVI-ETLRVSAGWMLLRSGVDLLVY-----LL 153
 ||| : ||| : ||| :
D8 26 LOGARAVPPSPPAACNLITVNTAVNQTLVNSASVLI-SVAPSVAVNASCLQSARGAL 83
 ||| : ||| : ||| :
QY 154 AHGALXILPSPSCAYQVCGSPLYQ-----IC-----ATTDIWP 186
 ||| : ||| : ||| :
D8 84 GSCITDLGRTSGCCCTQRCTALHNLTDCVNGLVRAVCMGSDGNATVNTTARAHTPLF 143
 ||| : ||| : ||| :
QY 187 SVSASYRPTPVGRN---FTNLRFAQIQQSSSRQEAQPKLALPSRGTKRHLSTSTSVPS 243
 ||| : ||| : ||| :
D8 144 SVYNNCIDTPAPYNCSTFNVTSVIPTSPSPSPSPSPSPSPSKASPSPSKASPS 203
 ||| : ||| : ||| :
QY 244 -AKKARCYPVRVEEGPHRVQLPTPGKK-----SWVSPARSPEVPTA 285
 ||| : ||| : ||| :
D8 204 PSPKASPSPKASPAPSPQSPSTPSKASPVASPOOSPTPSRPSTPSPTSPTS- PKA 262
 ||| : ||| : ||| :
QY 286 EKDLSSKGKVSDLISLGVCCKHKPSTLSLP---PRQAFOLRPFIEIRHFYLSRGDG 342
 ||| : ||| : ||| :
D8 263 SPFPSPSPSPSP-SLSPKVPSPSTPTGSPAASPGSGFRAS-----PPGGG 306
 ||| : ||| : ||| :
QY 343 QERLNPSLELLNLPNTTGARRLVEIIFLGSRRPTSCP 380
 ||| : ||| : ||| :
D8 307 PPMSP-----RLSEPTIPPTRSPIPIPSFIPTSP 339
 ||| : ||| : ||| :

RESULT 11

S49627
regulatory protein ARG2 - yeast (*Saccharomyces cerevisiae*)
N:Alternate names: protein YML099c
C:Species: *Saccharomyces cerevisiae*
C>Date: 02-Dec-1994 #sequence_revision 10-Feb-1995 #text_change 12-Dec-1997
C:Accession: S49627; A25064
R:Gentiles, S.; Bowman, S.
submitted to the EMBL Data Library, November 1994
A:Reference number: S49627
A:Accession: S49627
A:Molecule type: DNA

365 -GKEVLEKARGSTYGTTPRPVPPKPRPEVPS-DETAATSHGS-----AQVPEPTIHLAA 416
155 HCALYLL-----VPPSCAYQVCGSPLOYICATTDIWPSVSASYR-----PTRPVGRN 201
417 QGMAYPLHEHQGNAPCPVAQAAPTLPFV-SPGDLPGVFDGVRVACAPVAPAGPIVRP 475
202 FTNLRFLQOIKSSSRQEPAPKPLALPSRGTKRHLSTSTSVPSAKKAR-CYPVPRVEEGPH 260
476 W-----EPSLTQAAGQAFV-PVR--POHMEVEVPVPTVALERPVVPKP-VRPAPP 522
261 RQVLPTPSGKS-----WVPSPARSPVPTAEKDLSSGKGVSDLSLSGSV--- 305
523 KIAMQGPGETSGIRARERWRPAPWTPNPPRSP-----SQMSYRDLRLARAEQAQVKQA 576
305 CCKHKPSSTLSP-----PRQNAFQLRPFTIETHFLYRSGDGOERLNPSEL-LSN 354
577 SVEVQPPOLTQVSQOPMEGPLVPEQQMFGAPFSQVADV- -RAPGVPAHQPOYFDLPL 634
355 LQPNLTGA 362
635 IOPISQGA 642
RESULT 9
T00356
hypothetical protein KIAA0690 - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 01-Feb-1999
C:Accession: T00356
R:Rishikawa, K.; Nagase, T.; Suyama, M.; Miyajima, N.; Tanaka, A.; Kotani, H.; Nomura, N.
DNA Res. 5, 169-176, 1998
A:Title: Prediction of the coding sequences of unidentified human genes. X. The complete
A:Reference number: Z14142
A:Accession: T00356
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1214 <ISH>
A:Cross-references: EMBL:AB014590; NID:d1204345; PID:d1032626
A:Experimental source: brain; clone HK03594
C:Genetics:
A:Note: KIAA0690
Query Match 1.9%; Score 109.5; DB 2; Length 1214;
Best Local Similarity 21.3%; Pred. No. 5.2;
Matches 121; Conservative 78; Mismatches 191; Indels 179; Gaps 31;
QY 23 PLATFVRRLGPEGRRLLVQPG---DPKIVRTLVACQLVCM-HWGSOPPPADLSFHVQS--- 76
DB 508 PLA---NTLKSAMDLAQAGSTVESKIYDTLOWQMTLLPGFCTRPDTVAISFKGLARTL 564
QY 76 -----SLKELVARVVQQLCERNE-----RNVLAFGFELLNE--ARG--- 110
565 GMAISERPDRLRVTVQALRTLTKCCQAEADRAEVSREAKNPLPILFNLYGQPVAAQTDP 624
QY 110 GPPMAFTSSVRSYLPNTVIETLRVSGAWMLLSRYGDDLL-----VYLLAHCALYLLVP 163
DB 625 APPRAVLETIRYL--TITDOLVNS---LLEKASEKVLDPASSDFTPLSLVDLVVALA 678
QY 164 PCSAYQVCGSPLOYICATTDIWPSVSASRYTPRPVGRNFTNLRFLQOIKSSSRQEPAPKPL 223
DB 679 P-CADEAAISKLY-----STIRPYLESKAHGQVKKAYRVLEEVCS--POGPGAL 725
QY 224 ALPS--RGTNRHL--SLTSTSVPSAK-KARC-YPVPRVEEGPHROVLPTPSGKSWVPSA 277
DB 726 FVQSHLEDLKKTLDSLRSTSPAKRPLKCLLHIYRKVLSAEHKEFI-----IA 774
QY 278 RSEPVTAEDKSSKGVSDLSLSGVCCKHKPSSTLLSPPRQNAFQLRPFTIETHFLY 337
DB 775 LIPEVILCTKEVSQA-----RKNAFAL--LVEMGHAF 806
QY 338 SRGDGOERLNPSSLNLQNLNLTGARKRLVFIIFLGRPRTSGPLCRTHRLSRRYWOMRPL 397

A:Residues: 1-880 <GEN>
 A:Cross-references: EMBL:Z46660; NID:9575702; PID:9575703; MIPS:YML099c
 R:Messequy, F.; Dubois, E.; Descamps, F.
 Eur. J. Biochem. 157, 77-81, 1986
 A:Title: Nucleotide sequence of the ARGRII regulatory gene and amino acid sequence homol
 A:Reference number: A25064; MUID:86220196
 A:Accession: A25064
 A:Molecule type: DNA
 C:Genetics: 1-3, 'P', 5-128, 130-131, 'L', 132-282, 'V', 284-344, 'V', 346-365, 'Q', 367-548, 'A', 55
 A:Residues: 1-3, 'P', 5-128, 130-131, 'L', 132-282, 'V', 284-344, 'V', 346-365, 'Q', 367-548, 'A', 55
 A:Gene: SGD:ARG81; ARG82
 A:Cross-references: MIPS:YML099c; SGD:S0004565
 A:Map position: 13L
 C:Superfamily: unassigned GAL4-type zinc cluster proteins; GAL4 zinc binuclear cluster h
 C:Keywords: DNA binding; nucleus
 F:16-53/Domain: GAL4 zinc binuclear cluster homology <GAL4>

Query Match 1.8%; Score 108; DB 2; Length 880;
 Best Local Similarity 20.7%; Pred. No. 4.2;
 Matches 98; Conservative 69; Mismatches 165; Indels 142; Gaps 26;
 QY 93 ERNVLAFFGELLNEARGPPMAFTSSVRSYLPNTVITLRSVGMMLLSRVGDDLLVYL 152
 Db 469 QKFLGGLGIELRQASNF--LRCLNTKSSIPEKYKDLTA-----ILSMNSIDVWGT 520
 QY 153 LAHCALYLLVPPSCAYQVCGSPLYQIC-----ATTDIMPSVSASYPTRPVGRNFTNLR 207
 Db 521 MADCODHL-----ALCEDFVESRMKLRPNIS-----EKTKLHRIFSFLKL 561
 QY 208 LQOIKSSRQEQAPKPLAPSRGTRKHLSL--TSTSPSAAKARCYVPVVEGPHROVLEPT 266
 Db 562 IDOSTALDVKRAKEIILPSEEDDNYKPLDTSNATSSSEPR---VDVVQEGFLREALNE 618
 QY 267 PSKG---SWVPSPARSPEVPTAEKDLSSGKVDLSGSCCKKPKSPSTSLSPPRONA 323
 Db 619 NGDKTHIEVKEP-----ITNVSADST-----PST--Tpp----- 648
 QY 324 FOLRFETRHLYSRGQERLNSFLSLNLPNLTGARRLV-----EIFFLGRS----- 375
 Db 648 --IFTNIATESY-YNKSD-----ISKLVKTDENIIGTDSLYGLPNSLILFFSDCVRI 698
 QY 375 -----PTSGPLCETHLSRRYQWMLPFOQ-----LLVNHA 407
 Db 699 RHNEYNYLTYPVKRNFELNFKLLKWKSEWNEHQNSEGSKSFNSTAEALYHHTM 758
 QY 408 QOYVRL-----RS-HCRF-RTANQOQVTDALNTSPHMLDLRLHSSP--WQVYGLR 456
 Db 759 SFYSLIIYFTMARSLNCQFLQNVAKVLDHLN-AMEELVDQKKVIVPLIWO--GFMA 815
 QY 457 ACLKVVASLWGTNRHRRERFKNLKFIS---LGKY-GKLSQELMKMKVED 506
 Db 816 GCAC-----TDENRQOEFRRAAKLAESGVGSYGARQVMLEVWRRKED 860

RESULT 12

T01397
 LTR transposon polyprotein homolog T419.16 - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 24-Mar-1999
 C:Accession: T01397
 R:Parnell, L.D.; Gnoj, L.; de la Bastide, M.; Hameed, A.; Habermann, K.; Schutz, K.; Hua
 submitted to the EMBL Data Library, May 1998
 A:Description: Genomic sequence of BAC T419 from Arabidopsis thaliana, chromosome IV, ne
 A:Reference number: Z14314
 A:Accession: T01397
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-1456 <PAR>
 A:Cross-references: EMBL:AF069442; NID:93242970; PID:93924509
 A:Experimental source: cultivar Columbia
 C:Genetics:
 A:Map position: 4

A:Introns: 129/1
 A:Note: T419.16

Query Match 1.8%; Score 107; DB 2; Length 1456;
 Best Local Similarity 19.7%; Pred. No. 11;
 Matches 169; Conservative 108; Mismatches 288; Indels 294; Gaps 46;
 QY 88 LCERNERNVLAFFGELLNEA---RGPPMAFTSSVRSYLPNTV----- 128
 Db 603 LSEKRRHIVEMGLTLLSHASVPKTYWYAF--SVAYILNRLPTLLOLQSPQKLFQ 660
 QY 128 ---JETLRVSGA---WMLLSR--VGDD-----LLVYLLAHCA-LYLLVPPSCAYQVCG 172
 Db 661 PPNEYKLVKFCACYPWLRYNRHKLKEDSKQCAFMYSLTQSAYLCLHTCRLTYSRH 720
 QY 173 SPLYQIC---ATTDIMPSVSASYPTRP-----VGRNFTNLR 208
 Db 721 VQDFERCFFETTFNGVSTSQEQRSDSAPWPSHTTLPPLVLPAPPCGLGPHL----- 775
 QY 209 QOIKSSRQEQAPKPLAPLS-----RGTKRHLSTSTSVPSAKK--ARCYPVPRVEGPHR 261
 Db 775 -----DTSPPRPPSPPLCTTOYSSSNLPSISSSPSSSEPTAPSHNGPQTAQPHQ 826
 QY 262 ---QVLTPSGKSWVP-----SPARSPEVPTAEKDLSSGKVDLSL 302
 Db 827 TQNSNSPILNNPNPNSPSPNONSPLQSPISPHIPTSTSISE----- 876
 QY 303 SVCKKHKPSTSLSPRONAFOLRPFETR-----HFLYSRG-DGOERLNPFS-LLS 353
 Db 876 -----PNSPSSSTSTPPLPVLPAPPPIQVNAQAPVNTHSMATRAKDGIRKKNQKYSAT 931
 QY 354 NLQPNLTGARRLVEIFLGRSRTSGPLCETHLSRRYQWMLPFOQLLVNHAECQYVRL 413
 Db 932 SLAAN-----SEPTAIQAMKDDR-----WQAMGSEINAAQ----- 963
 QY 414 LRSHCFRTANOQVTDALNTSPH--LMDLLRLHSSPQVYGFELRACLCKVYVASLWCTR 471
 Db 963 -----IGNHTWDLVPPPPSVTIIVGCRWIFTKKNSDGLNRYKARLVAGK----- 1009
 QY 472 HNERREFKNLKFISLGVKGLSL-----QELMKMKVEDCH-----WLRSSP 514
 Db 1009 YNORGLDYAETFSPIVKSTSIKIVLGVAVDRSWPIROLDVNNAFLOGLTDEVIMSOP 1068
 QY 515 G---KDRVPAAEHLRERI-----LATFLW-----LMDT--YVYQLLRSP 550
 Db 1069 GFVDKDR-PDYVCLRKALYGLKQAPRAYVELTYLLTVGFVNSISDTSFLVLRGSI 1127
 QY 551 F-----YITESTFQKRLFFYKKSWSKLSQISGRQH-----LERVRLRE---LSQE- 595
 Db 1128 IYMLVYVDDILITGNDTVLLKHTLDALSQRFSVKEHEDLHVFLGIEAKRVPQGLHSQR 1187
 QY 595 ---EVRHHQDTWLAMPIC-----RLRF-----IPKNGLRPIV-NWYSMGT 634
 Db 1188 YTLDLARTNMLTAKPVATPMATSPKTLHSGTKLPDTEYRGIVGSLQYLAFTRPDLSY 1247
 QY 634 ALGRKQKQHQETORLKTFLSMLNYERTKHPHLMGSSVGLMNDIYRTWRAFLVRVRLDOT 693
 Db 1248 AVNRLSQYMH-----MPTDD---HWNALKRVLRYLAGT 1277
 QY 694 P-RMYFVKADVT---GAY-DAIPOGLVEVVAN-----MIRHSESTYCIQYAVVRRDS 742
 Db 1278 PDHGIFLKKGNTLSLHAYSADNAGDDDDYVSTNGVYILVGLGHPISWSSKKQGVVRSST 1337
 QY 743 QGVHKSFRFRQVTTLSDLQ 761
 Db 1338 EAE---YRSVANTSSELQ 1352

RESULT 13

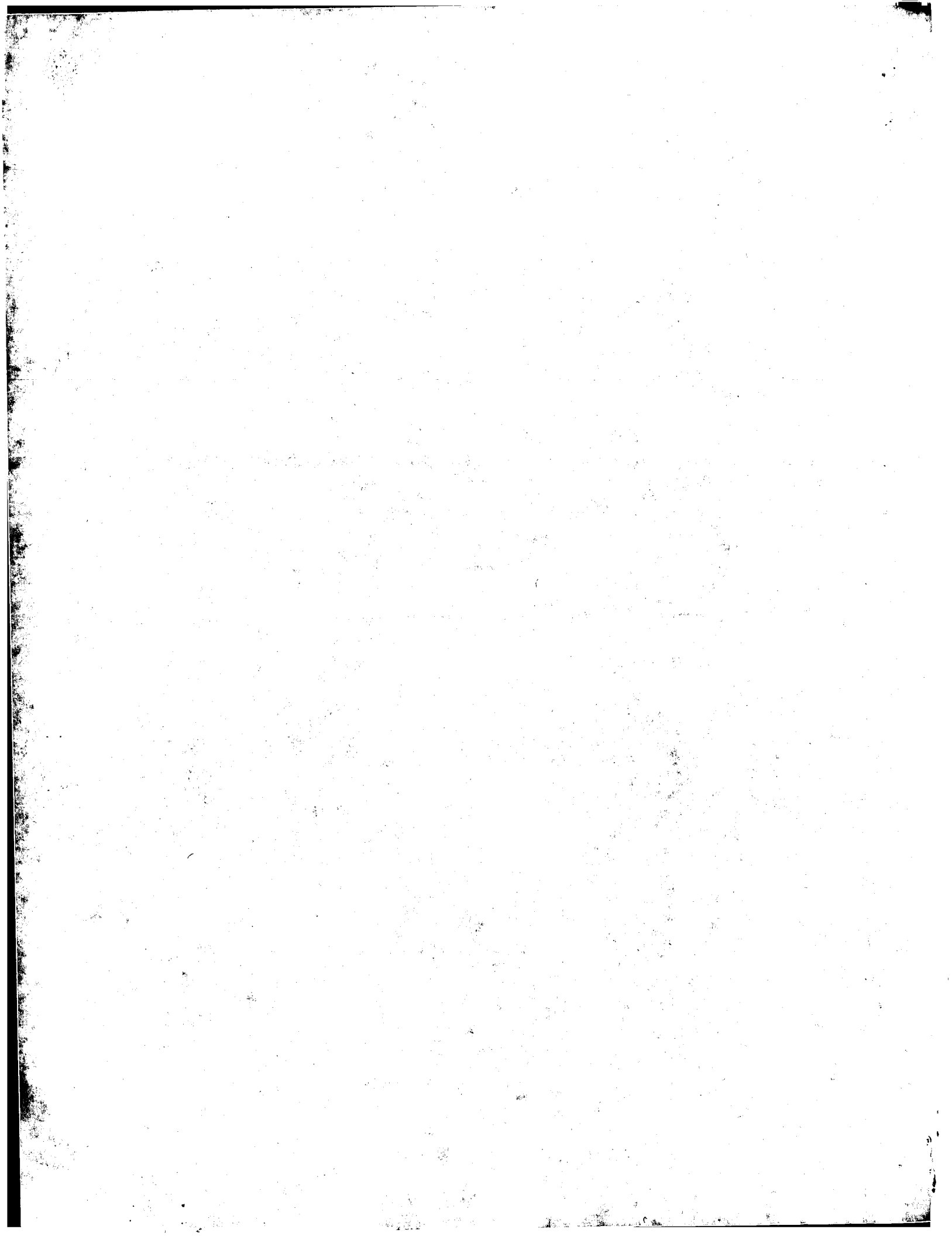
C71436
 Probable resistance gene - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse ear cress)

QY 803 LHFLRHSVVKIGDRYCTOCQIQPOGSSLSLTLCSLCFGDMENKLFPAEVQRDGLLLREVDD 862
Db 831 -----LILADDNFTIVEAIKEGRCI-----YNNMK-----AFIRY--- 862
QY 853 FLLVTPHLDQAKTFLSTLVHGVPEYGCINLQKTVNFPVE--PGLTGGAPYQLPAHCL 920
Db 862 --LISSNIGEVASIFITAILGIPD--SLAPVOLLVNLVTDGLPATALGNP---PEHDV 914
QY 921 FPMC-----GILL-----DTQLEVFCDYSG 941
Db 915 MK-CKPRHRNDNLINGLTLRLYIVIGTVGIATVSIYIYWMFVDPMDNHTLINFYQLSH 973
QY 942 YAOISIKTSLTFQ-----SVFKAGKTMKNKL-LSVLRL-----KCHG 977
Db 974 YNQ--CKTWSFNVNKYDMSDELCSYFSAGVKVAKSTLSLSVLVLIEMFNALNALSEYNS 1031
QY 978 LFLDLQVNSLQTVCIINIYKIFLQAYRFHACVIQLP 1013
Db 1032 LFLPPWRNMYLVLATIGSLFL-----HCLIIYFP 1061

RESULT 15

S40013
hypothetical protein 2 - Calothrix sp. (PCC 7601)
C:Species: Calothrix sp.
A:Variety: PCC 7601
C:Date: 08-Jun-1994 #sequence_revision 12-Apr-1996 #text_change 17-Mar-1999
C:Accession: S40013; S35079
R:Ferat, J.
submitted to the EMBL Data Library, April 1993
A:Reference number: S40013
A:Accession: S40013
A:Molecule type: DNA
A:Residues: 1-584 <FE1>
A:Cross-references: EMBL:X71404; NID:g396226; PID:g396228
A:Experimental source: PCC 7601
R:Ferat, J.L.; Michel, F.
Nature 364, 358-361, 1993
A:Title: Group II self-splicing introns in bacteria.
A:Reference number: S35078; MUID:93323980
A:Accession: S35079
A:Molecule type: DNA
A:Residues: 224-338:513-565 <FE2>
A:Cross-references: EMBL:X71404
A:Experimental source: PCC 7601

Query Match 1.8%; Score 106; DB 2; Length 584;
Best Local Similarity 22.5%, Pred. No. 3.2;
Matches 62; Conservative 33; Mismatches 111; Indels 70; Gaps 11;
QY 676 IYRTWRAFLVRALDQTPRMFYKADVTGAYDAIPQGLVEYVANMIRHSESTYCIROY 735
Db 50 LMRWSGKALAVRKVTO-DNQGKKAAGIDGVKSLKPSARLTLVNMKLNHK-----VKAT 103
QY 736 AVVRDSOGVHKSFRRQVTTLSL-----LQPYMGQFLKHLQSDA- 777
Db 104 RRVVTPKGVNEKR-PLGIPTMQDRATQSLVKALEPEWEAKFEPNSYGFPGRNAHDAR 162
QY 777 SALRNSVVEQISISMNESSSLFDFFLH-----FLRHSVVKIGDRC 817
Db 163 EAFINSIRYSNKKVLDADISKCFKINHEKLLTINTFTMRROIKAWLKAGVLDNGHFS 222
QY 818 YTCGGIPOGSSLSLTLCSLCFGDMEN--KLFAEVQRD-----LLRFVDDFLVLT 867
Db 223 ET-TEGTPQGGVISPILLANIALHGLEKLVKFEFAASQGGKVNQNSISLIRYADDFVILA 281
QY 868 PHLDQ-----AKTFLSTLVHGVPEYGCINLQKT 896
Db 282 PNKTQIIIVLKEIVKTLA-----EMGLELNPNT 310



GenCore version 4.5
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 14, 1999, 04:51:16 ; Search time 12.16 Seconds
(without alignments)
2608.308 Million cell updates/sec

Title: US-09-042-460-2
 Perfect score: 5901
 Sequence: 1 MTRAPRCPAYRSLRSRYRE.....TILKAADPALSTDFQILD 1122
 Scoring table: BLOSUN62

Searched: 77977 seqs, 28268293 residues
Database : SwissProt_37.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	124	2.1	4540	1	DYHC_PARTE	Q27171 paramecium
2	111.5	1.9	1415	1	ICP4_EBVMG	Q02362 marek's dis
3	110	1.9	812	1	EBN3_EBV	P12977 epstein-bar
4	108	1.8	880	1	ARG2_YEAST	P05085 saccharomyce
5	105	1.8	1493	1	MEK1_RAT	Q42935 rattus norv
6	104.5	1.8	916	1	RTJK_DROFU	P13329 drosophila
7	103.5	1.8	719	1	DEND_RAT	P30617 rattus norv
8	102.5	1.7	758	1	CY14_NEUCR	P23622 neurospora
9	101	1.7	677	1	T2D5_HUMAN	P49848 homo sapien
10	99.5	1.7	810	1	ADMD_YEAST	P15274 saccharomyce
11	98.5	1.7	2163	1	BRR2_YEAST	P32639 saccharomyce
12	98	1.7	745	1	CUL2_HUMAN	Q13617 homo sapien
13	98	1.7	1239	1	NME3_MOUSE	Q01098 mus musculu
14	97.5	1.7	916	1	RTJK_DROME	P21328 drosophila
15	97	1.6	3148	1	HD_FUGRU	P51112 fugu rubrip
16	97	1.6	1174	1	KCRF_STRPU	P18294 strongyloce
17	97	1.6	1314	1	SS22_YEAST	P25390 saccharomyce
18	96.5	1.6	585	1	ASN1_LOTJA	P49092 lotus japon
19	96.5	1.6	446	1	TBB5_CHICK	P09653 gallus gall
20	96.5	1.6	856	1	UN51_CAEEL	Q23023 caenorhabdi
21	96.5	1.6	1336	1	VEGR_RAT	P35767 rattus norv
22	96.5	1.6	2870	1	YAOS_SCHPO	Q10105 schizosacch
23	96	1.6	480	1	P53_HORSE	P79892 equus cabal
24	96	1.6	1155	1	ROM1_YEAST	P53046 saccharomyce
25	96	1.6	2156	1	RRPL_PUUMH	P27176 puumala vir
26	96	1.6	320	1	RT86_ECOLI	P23070 escherichia
27	96	1.6	976	1	SOPL_HUMAN	Q15431 homo sapien
28	96	1.6	865	1	YMW6_ECOLI	P06612 escherichia
29	96	1.6	1277	1	YPM6_YEAST	Q04264 saccharomyce
30	95.5	1.6	714	1	CLPB_MYCCE	P47597 mycoplasma
31	95.5	1.6	606	1	KRAA_HUMAN	P10398 homo sapien
32	95.5	1.6	1233	1	NME3_HUMAN	Q14957 homo sapien
33	95	1.6	2554	1	7LES_DROME	P13368 drosophila
34	95	1.6	1093	1	AF17_HUMAN	P55198 homo sapien
35	95	1.6	1547	1	RRO_PMW	P20951 papaya mosa
36	94.5	1.6	1260	1	LINI_NYCCO	P08548 nycticebus
37	94.5	1.6	1660	1	VIT6_OSCBR	Q94637 oschelus br
38	94.5	1.6	1118	1	Y1P1_YEAST	P40456 saccharomyce
39	94.5	1.6	2717	1	ZEP1_HUMAN	P15822 homo sapien
40	94	1.6	1182	1	ABL2_HUMAN	P42684 homo sapien
41	94	1.6	2090	1	N214_HUMAN	P35658 homo sapien
42	94	1.6	1291	1	P1P4_BOVIN	P08487 bos taurus
43	94	1.6	985	1	REV1_YEAST	P12689 saccharomyce

44	94	1.6	445	1	TBB_PLAFK	P14643	plasmodium
45	94	1.6	656	1	UL25_HCMVA	P16761	human cvtom

ALIGNMENTS

```

RESULT 1
DYHC_PARTE
ID ID DYHC_PARTE STANDARD; PRT; 4540 AA.
AC Q27171;
DT DT 01-NOV-1997 (REL. 35, CREATED)
DT DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE DE DYNEIN HEAVY CHAIN, CYTOSOLIC (DYHC).
GN GN DHC-8.
OS OS PARAMECIUM TETRAURELIA.
EUKARYOTA; ALVEOLATA; CILIOPHORA; NASSOPHOREA; PENICULIDA; PARAMECIUM.
[1]
RN RN SEQUENCE FROM N. A.
RP RP STRAIN-STOCK 51;
RC RC MEDLINE; 96157890.
RA RA KANDL K.A., FORNEY J.D., ASAI D.J.;
RT RT "The dynein genes of Paramecium tetraurelia: the structure and
RL RL expression of the ciliary beta and cytoplasmic heavy chains.";
RM MOL. BIOL. CELL 6:1549-1562(1995).
CC CC -1- FUNCTION: DYNEIN HAS ATPASE ACTIVITY. CYTOPLASMIC DYNEIN ACTS AS A
CC CC MOTOR FOR THE INTRACELLULAR RETROGRADE MOTILITY OF VESICLES AND
CC CC ORGANELLES ALONG MICROTUBULES.
CC CC -1- SUBUNIT: CONSISTS OF AT LEAST TWO HEAVY CHAINS AND A NUMBER OF
CC CC INTERMEDIATE AND LIGHT CHAINS.
CC CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC CC -1- SIMILARITY: BELONGS TO THE DYNEIN HEAVY CHAIN FAMILY.
-----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-sib.ch/
CC CC or send an email to license@isb-sib.ch).

```

Query Match 2.18; Score 124; DB 1; Length 4540;

Matches 127; Conservative 111; Mismatches 214; Indels 234; Gaps 333; Sequence Similarity 10.5%; FIED. NO. 2.3;

[illegible]

Query Match	1.9%;	Score 111.5;	DB 1;	Length 1415;
Best Local Similarity	21.4%;	Pred No. 3.6;		
Matches	97;	Conservative	50;	Mismatches 129; Indels 177; Gaps 25;
DOMAIN	155	200	SER/PRO-RICH.	
SEQUENCE	1415 AA;	154936 MW;	21FB00A6 CRC32;	
QY	172	GSPLYQICATTDIWPVSASYSRTPRVGRNFTNLRFLOQIKSSRQ	----	EAPKP----- 223
DB	52	GQPEPPLVPVD-KNNFSTDWRSP	-----	RSGPKKDFCGDLPAFLTSGP 96
QY	223	-IALPSRGTKKHLSTSTVPSKAKKARYPVVEGPHRQVLPTPSGKSWVPSPARSPE	281	
DB	97	RLTTPSSGRMSELPHTTSS	-----	PRSSPRPR--GPE-----TPSNEHIISPPRNP 143
QY	282	VPTAKDLSSGKYS-DLSLSGSCCKKHKPSSTLL	-----	SPPRQNA-FQL 326
DB	144	SNTHRNV---CHVSRSPPSSSSSSSSSSSLIVLSPSSSRSPSPPRPRADSS	200	
QY	327	RPFIETRHLYSGDQGRNLNPSLNLNPNLTGARRLVEIFL	----	GSRPT 377
DB	201	RP-----RRGRSGNRGSRGSPQSKGASPRTKLEDEDYLPQETANRRGGPRG	251	
QY	378	SGPLCETHLSRR	-----	YWQMRPL-----FOOL 401
DB	252	RP--KSGRAVORNDIQVTSSGLADTSPYDLGCVWWEV-PLPPGRCWFGGLGHRQA	308	
QY	402	LYNHAEQCVRLLRSHCFRTANQOVT	-----	DALNTSPHMLDLRLHSSP 448
DB	309	LTDSP-----IVEAIHRENTSHGVPVVEEMKYAKQDALVNLFH-KSMKYVNLN	361	
QY	449	WQ-----	-----	YGFELRACLCKVWSAS-----LWQTRH- 473
DB	362	WHHGLKSPADALNHIYVQFQSDSPGAAVTGTVNRCPHIAKMRKLLWAFPHI	421	
QY	473	-----NERFFNKKFFI--SLGK-YGKLSQE	497	
DB	422	AASIAMTRYCKDQKTFELRSKKAYASMAFPD	454	
RESULT	3			
EBN3_EBV		STANDARD:	PRT:	812 AA.
ID	EBN3_EBV			
AC	P12977	Q66540;		
DT	01-OCT-1989	(REL. 12, CREATED)		
DT	01-OCT-1989	(REL. 12, LAST SEQUENCE UPDATE)		
DT	01-NOV-1997	(REL. 35, LAST ANNOTATION UPDATE)		
DE	EBNA-3	NUCLEAR PROTEIN (EBNA-3A).		
GN	BRF1.			
OS	EPSTEIN-BARR VIRUS (STRAIN B95-8) (HUMAN HERPESVIRUS 4).			
OC	VIRUSES; DSDNA VIRUSES, NO RNA STAGE; HERPESVIRIDAE;			
OC	GAMMAHERPESVIRINAE; LYMPHOCRYPTOVIRUS.			
RN	[1]			
RN	SEQUENCE FROM N.A.			
RX	MEDLINE; 84270667.			
RA	BAER R.; BANKIER A.T.; BIGGIN M.D.; DEININGER P.L.; FARRELL P.J.;			
RA	GIBSON T.J.; HATFULL G.; HUDSON G.S.; SATCHWELL S.C.; SEGUIN C.;			
RA	TUFFNELL P.S.; BARRELL B.G.;			
RT	"DNA sequence and expression of the B95-8 Epstein-Barr virus genome.";			
RL	NATURE 310:207-211(1984).			
RN	[2]			
RP	CHARACTERIZATION.			
RX	MEDLINE; 86287322.			
RA	HENNESSY K.; WANG F.; BUSHMAN E.W.; KIEFF E.;			
PT	"Definitive identification of a member of the Epstein-Barr virus			
RT	nuclear protein 3 family.";			
RL	PROC. NATL. ACAD. SCI. U.S.A. 83:5693-5697(1986).			
RN	[3]			
RP	SUBCELLULAR LOCATION.			
RX	MEDLINE; 90266473.			

Mon Aug 16 10:21:27 1999

Db 759 SFYSLIIYFTMARSLNCFQNLQNVAKVLHNL-AMEELVDQKKVKIVPLIQ--GFMFA 815

QY 457 ACLCKVWSASLWGTNRHNRFFKNLKFFIS---LGKY-GKLSQLQELMMKMKVED 506

Db 816 GCAC-----TDENRQEFRRWAALAESGVGYNGARQVLMLEVRRRKED 860

RESULT 5

MEK1_RAT ID MEK1_RAT STANDARD; PRT; 1493 AA.
AC Q62925;
DT 01-NOV-1997 (REL. 35, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE MAPK/ERK KINASE 1 (EC 2.7.1.-) (MEK KINASE 1) (MEK 1)...
GN MEK1 OR MEK2.
OS RATTUS NORVEGICUS (RAT).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; RATTUS.
RN [1]
RP SEQUENCE FROM N.A., AND MUTAGENESIS.
RC TISSUE=BRAIN;
RX MEDLINE; 96224276.
RA VANDERBILT C.A., COBB M.H.;
RA "Cloning of rat MEK kinase 1 cDNA reveals an endogenous membrane-associated 195-kDa protein with a large regulatory domain."
RT PROC. NATL. ACAD. SCI. U.S.A. 93:5291-5295(1996)
CC -!- FUNCTION: CAN PHOSPHORYLATE AND ACTIVATE MAPKK 1 AND MAPKK 2 (MEK1/MEK2) WHICH LEADS TO PHOSPHORYLATION OF MAP KINASES. MOST POTENTIALLY ACTIVATES THE JNK/SAPK GROUP OF MAP KINASES, AND LESS EFFICIENTLY ERK2 OR P38.
CC -!- SUBCELLULAR LOCATION: MEMBRANE ASSOCIATED.
CC -!- TISSUE SPECIFICITY: MOST HIGHLY EXPRESSED IN SPLEEN, KIDNEY, AND LUNG.
CC -!- PTM: AUTOPHOSPHORYLATED.
CC -!- SIMILARITY: WITH THE CONSERVED CATALYTIC DOMAINS OF SER/THR-PROTEIN KINASES. BELONGS TO THE MAP KINASE KINASE FAMILY.
CC -!- SIMILARITY: CONTAINS 2 PH DOMAINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC
CC EMBL; U48596; G1354137; -
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PFAM; PF00069; Pkinase; 1.
DR HSSP; P24941; 1AQL.
KW TRANSFERASE; SERINE/THREONINE-PROTEIN KINASE; ATP-BINDING; PHOSPHORYLATION.
FT DOMAIN 2 5
FT DOMAIN 17 20 POLY-ALA.
FT DOMAIN 25 29 POLY-GLY.
FT DOMAIN 74 149 PRO-RICH.
FT DOMAIN 111 114 POLY-PRO.
FT DOMAIN 233 291 PRO-RICH.
FT DOMAIN 412 421 POLY-SER.
FT DOMAIN 439 555 PH.
FT DOMAIN 643 750 PH.
FT DOMAIN 1163 1168 POLY-GLU.
FT DOMAIN 1197 1200 POLY-ILE.
FT DOMAIN 1224 1489 PROTEIN KINASE.
FT DOMAIN 1230 1237 ATP (BY SIMILARITY).
FT BINDING 1253 1253 ATP (BY SIMILARITY).
FT ACT_SITE 1350 1350 BY SIMILARITY.
FT MUTAGEN 1369 1369 D->A: INACTIVATION.
SQ SEQUENCE 1493 AA; 161315 MW; 982785D9 CRC32;

Query Match 1.8%; Score 105; DB 1; Length 1493;
Best local similarity 18.5%; Pred. No. 12;
Matches 114; Conservative 79; Mismatches 209; Indels 214; Gaps 30;

QY 163 PPSYAYVCGSPLYQICATTDIWPVSASRYP---TRPVGRNFTNLRELQOI-KSSSROE 218
Db 112 PPPGAASRCGSHSAELAAARDGARSAPAGAEPPSAAPSGREMENKETLGLHKHMDRPE 171
QY 219 APKPLALPSRGTKRHLSTSTSVPSAKKARCPVPDVE----- 257
Db 172 -----ERMIREKLKATCPANKHEWLERENRRGPPVVKPIPIKGDG 212
QY 257 -----EGPHRQVLPTPSGKSWPSPARSPEVPTAEKDLSS---GKVSLSL-S 301
Db 213 SEMSNLAELQEGGAGSAAPAKGRR-SPSPGSSPSGRSGKPESGVRKRKRVPVFFOS 271
QY 302 GSVCCCKHKPSTSLSP--PRONAFOLRPFIEIETRYHLYSRGQOERLNP-SFLLSNLQPN 358
Db 272 GRITPPRRAPSPDGFSPYSPEETSRVNVKVRARLYLL-----QQIGPNSFLIGDSDP 325
QY 359 LTGARRLVEIIFLGSRPRTSGPLCRTHRLSRRYWQMRPLFQQLLVNHAECQYVRLLRSHC 418
Db 326 --NKYR---VFIGPQNCSCG-----RGTFCHLLFVMLRVFQLEPDPMLWRKTL 370
QY 419 RFRFANOQVTDALNTSPHMLDLLLHSSWQVYGFRLACLCCKVVSASLWGTNRHNRREF 478
Db 371 K-----NFEV-----ESLFQYHSRRSRRIKAPSRN----- 397
QY 479 KNLAKFTSLGKYGLSLQELQELMMKMKVEDCHLWLRSSPGKDRVPAAEHLRL---ERILATFL 535
Db 397 -TIQKFS-----RMSNCHTLSS--STSTSSSENSIKDEEQMCPICL 437
QY 536 FWMMDTYVYVOLLRSFFYITESTFOKNRLFFYRKSWSKLSQISGVQHLRVLRLSELEE 595
Db 438 LGMLEESLTVCEDCG-----RNKLHHHMSIWA-----EE 468
QY 596 VRHODTWLAMPICRLR-----IPKNGRLRPVNNMSYSN---GTRALGR 638
Db 469 CRNRREP-LICPLCRSKWRSHDFVSHELSSPVDSTSLRGVQVQSPSPQPVAGSQ---RR 524
QY 639 KQAQHF-----TQRLKTLFSLNLYERTKHPHL--MSSVLGMNDIYRTWPAFVLRVRL 690
Db 525 NQESNFNLTHYGTQIIPPA-----YKDLAEPWIOAFGMELVGCIL-FSRNNVREMAURLRL 578
QY 691 DOTPRMYFKADVTGA 706
Db 579 SH-----DVSGA 585

RESULT 6

RTJK_DROFU ID RTJK_DROFU STANDARD; PRT; 916 AA.
AC P21329;
DT 01-MAY-1991 (REL. 18, CREATED)
DT 01-MAY-1991 (REL. 18, LAST SEQUENCE UPDATE)
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
DE RNA-DIRECTED DNA POLYMERASE FROM MOBILE ELEMENT JOCKEY (EC 2.7.7.49)
GN POL.
OS DROSOPHILA FUNEBRIS (FRUIT FLY).
OC EUKARYOTA; METAZOA; ARTHROPODA; TRACHEATA; HEXAPODA; INSECTA;
OC PTERYGOTA; DIPTERA; BRACHYCERA; MUSCOMORPHA; EPHYDROIDEA;
OC DROSOPHILIDAE; DROSOPHILA.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1911.1;
RX MEDLINE; 91067678.
RA MIZROKHI L.J., MAZO A.M.;
RT "Evidence for horizontal transmission of the mobile element jockey between distant Drosophila species."
RT PROC. NATL. ACAD. SCI. U.S.A. 87:9216-9220(1990).
RL

us-09-042-460-2.rsp

Mon Aug 16 10:21:27 1999

```

RX MEDLINE: 94188926.
RA SANDAL N.N., MARCKER K.A.;
RT "Similarities between a soybean nodulin, Neurospora crassa sulphate
RL perasease II and a putative human tumour suppressor.";
RL TRENDS BIOCHEM. SCI. 19:19-19(1994).
CC -!- FUNCTION: UPTAKE OF SULFATE INTO THE CELL.
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -!- INDUCTION: HIGHLY EXPRESSED, BUT ONLY IN CELLS SUBJECT TO SULFUR
CC -!- LIMITATION, AND IT IS TURNED ON BY THE POSITIVE-ACTING CYS-3
CC SULFUR REGULATORY PROTEIN.
CC -!- SULFATE PERMEASE II IS MAINLY FOUND IN MYCELIA.
CC -!- SIMILARITY: BELONGS TO THE SULFATE PERMEASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: M59167; G168911; ALT_SEQ.
CC DR PIR: A37956; A37956.
CC DR PROSITE: PS01130; SULFATE_TRANSP: 1.
CC DR PFAM: PF00916; Sulfate_transp: 1.
CC DR TRANSPORT: TRANSMEMBRANE; GLYCOPROTEIN.
CC FT TRANSMEM 71 91 POTENTIAL.
CC FT TRANSMEM 103 123 POTENTIAL.
CC FT TRANSMEM 128 148 POTENTIAL.
CC FT TRANSMEM 171 191 POTENTIAL.
CC FT TRANSMEM 193 213 POTENTIAL.
CC FT TRANSMEM 271 291 POTENTIAL.
CC FT TRANSMEM 326 346 POTENTIAL.
CC FT TRANSMEM 363 383 POTENTIAL.
CC FT TRANSMEM 451 471 POTENTIAL.
CC FT TRANSMEM 474 494 POTENTIAL.
CC FT CARBOHYD 23 23 POTENTIAL.
CC FT CARBOHYD 578 578 POTENTIAL.
CC SEQUENCE 788 AA; 87864 MW; 4BB097A8 CRC32;

Query Match 1.7%; Score 102.5; DB 1; Length 788;
Best Local Similarity 21.6%; Pred. No. 7.2;
Matches 93; Conservative 52; Mismatches 170; Indels 115; Gaps 19;

QY 55 LVCWHGSGPPADLSFHOVSSLKE---LVARVVQRLCERNERNVLAFGFELNEARGG 110
DB 254 LVLVQMGKRYPROQRAFFVSTLRNVFIIILYLVSWLVNHRVDPKKAHKIL---GH 309
QY 111 PPMFTSSVRSYLPNTVIETLRVSGAMLLLSRVGDDLLVYLLAHCAL-----YLLV 162
DB 310 VPSGQHGKAPRLDNEILSA--ISG-----DIPTTILVLLIEHIAISKSFGRVNNYII 360
QY 163 PPSGAYQVCG-----SPLXQI-----CATTDIWPVSASVSRTRPVGR 200
DB 361 NPSQELVAIGFTNLGPFGLALPRYRIILENHOGQSCSHASRW----HYLRPRPARS 416
QY 201 NFTNRLFLQKSSRSQKAPLAPLSRGTRKHLSTSTSVPSAKKARCPV-----253
DB 417 LCAHLRLLLHPQOORRRH-DHPRRRDPDSSKGVYKFWLTS-----PLEVVIFAG 465
QY 253 -----PRVEGPHQVLPPTSGKSW--VPSPAR---SPEVPTAKDLSSKGKVDLSLS 301
DB 466 VEVSIFTSIENGIYVTAASGAVLLWRIAQSPGKFLGTOTEIYAPRELVRGSK--DSGRS 523
QY 302 GSVCKKHKPSSTSLSPRONAFQLRPFETTRH---FLYSGDGOERLNPFSLLSNLPQN 358
DB 524 RVACCKSR-STTAFSLDRDDLNLNQLSTPWFIFVYRGEGLNYVNSAKHLDNLFIH 582
QY 359 L--TGARRLVEIIFLSRP-----RTSGPLCRT-----HRLSR 390
DB 583 VFKHTRTELNFELKGLDRPNDPGPRPRLNRRARFAPDPAHHPRLLRQLRRCDDR 642

```

```

QY 391 YWMRPLFOQ 400
DB 643 PRLQDLRNQ 652
...
RESULT 9
T2D5_HUMAN STANDARD; PRT; 677 AA.
ID T2D5_HUMAN
AC P49848;
DT 01-OCT-1996 (REL. 34, CREATED)
DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE TRANSCRIPTION INITIATION FACTOR TFIIID 70 KD SUBUNIT (TAFII-70)
DE (TAFII-80) (TAFII80).
GN TAF2E OR TAFII70.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE: 94085406.
RA WEINZIERL R.O., RUPPERT S., DYNLACHT B.D., TANESE N., TJIAN R.;
RT "Cloning and expression of Drosophila TAFII60 and human TAFII70
RT reveal conserved interactions with other subunits of TFIIID.";
RL EMBL J. 12:5303-5309(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE-PLACENTA;
RX MEDLINE: 95396764.
RA HISATAKE K., OHTA T., TAKADA R., GUERMAH M., HORIKOSHI M.,
RA NAKATANI Y., ROEDER R.G.;
RT "Evolutionary conservation of human
RT TATA-binding-polyptide-associated factors TAFII31 and TAFII80 and
RT interactions of TAFII80 with other TAFs and with general
RT transcription factors.";
RL PROC. NATL. ACAD. SCI. U.S.A. 92:8195-8199(1995).
CC -!- FUNCTION: TFIIID IS A MULTIMERIC PROTEIN COMPLEX THAT PLAYS A
CC CENTRAL ROLE IN MEDIATING PROMOTER RESPONSES TO VARIOUS ACTIVATORS
CC AND REPRESSORS. BINDS TIGHTLY TO TAFII-250 AND ALSO DIRECTLY
CC INTERACTS WITH TAFII-40.
CC -!- SUBUNIT: TFIIID IS COMPOSED OF TATA BINDING PROTEIN (TBP) AND A
CC NUMBER OF TBP-ASSOCIATED FACTORS (TAFS).
CC -!- SUBCELLULAR LOCATION: NUCLEAR.
CC -!- SIMILARITY: BELONGS TO THE TAF2E FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: L25444; G437385;
CC DR EMBL: U31659; G1136306;
CC DR TRANSFAC: T00783;
CC KW TRANSCRIPTION REGULATION; NUCLEAR PROTEIN.
CC SEQUENCE 677 AA; 72668 MW; C270C85F CRC32;

Query Match 1.7%; Score 101; DB 1; Length 677;
Best Local Similarity 21.0%; Pred. No. 7.5;
Matches 89; Conservative 32; Mismatches 114; Indels 188; Gaps 20;

QY 7 CPAVRSLSLRSYREVWPLATFVRRLGPEGRRLVQDPKIVRTLVLAQCLVCMH-----60
DB 311 CIVSRQLCLRPDNDVNHWRDFAAR-----LVQA--ICKHESTTN 349
QY 60 -----WGSOPPPADLSFHQVSSLSKELVARVVO-----RLCERNERNVLAFG 101
DB 350 NIOSRITKTKFTKSWDEKTPWTRTRYSGIAGLAELGHVDIKTLILPRLQOEGE-403
QY 102 ELLNEARGGPPMAFTSSVRSYLPNTVIETLRVSGAMLLLSRVG--DOLLVYLLAHCA--L 158

```


RA HYMAN R., KAYSER A., KOMP C., LASHKARI D., LEW H., LIN D.,
RA MOSDALE D., NAKAHARA K., NAMATH A., NORGEN R., OEFNER P., OH C.,
RA PETEL F.X., ROBERTS D., SEHL P., SCHRAMM S., SHOGREN T., SMITH V.,
RA TAYLOR P., WEI Y., YELTON M., BORSTEIN D., DAVIS R.W.,
RL SUBMITTED (DEC-1994) TO EMBL/GENBANK/DDBJ DATA BANKS.
RN [2]
RP SEQUENCE OF 1-169 FROM N.A.
RA MULLIGAN J.T., DIETRICH F.S., HENNESSEY K.M., SEHL P., KOMP C.,
RA WEI Y., TAYLOR P., NAKAHARA K., ROBERTS D., DAVIS R.W.,
RL SUBMITTED (FEB-1993) TO EMBL/GENBANK/DDBJ DATA BANKS.
RN [3]
RP CHARACTERIZATION.
RC STRAIN=BDY473;
RX MEDLINE; 96304576.
RA NOBLE S.M., GUTHRIE C.,
RT "Identification of novel genes required for yeast pre-mRNA splicing
by means of cold-sensitive mutations.";
RL GENETICS 143:67-80(1996).
CC -!- FUNCTION: REQUIRED FOR PRE-MRNA SPLICING.
CC -!- SUBCELLULAR LOCATION: NUCLEAR (POTENTIAL).
CC -!- SIMILARITY: BELONGS TO THE SKI2 SUBFAMILY OF HELICASES.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U18922; G603413;
CC PIR; S30856; S30856.
CC SGD; L0003100; BREAD; 1.
CC PFAM; PF00271; helicase_C; 1.
CC KW HELICASE; ATP-BINDING; NUCLEAR PROTEIN; MRNA PROCESSING.
CC FT NP_BIND 79 86 ATP (POTENTIAL).
CC FT NP_BIND 521 528 ATP (POTENTIAL).
CC FT SITE 634 637 DEIX BOX.
CC FT SEQUENCE 2163 AA; 246183 MW; 48DB127B CRC32;
CC
CC Query Match 1.7%; Score 98.5; DB 1; Length 2163;
CC Best Local Similarity 18.0%; Pred. No. 63;
CC Matches 171; Conservative 117; Mismatches 317; Indels 345; Gaps 46;
QY 168 YQVC-----GSP---LYQIC--ATDIMPVSASYRTPRPGVGRNFTNLRFLQIKSSRQ 217
DB 1113 FEICLRKGWGHPTRMLLNLCKSATKMW-----TNCPL-LRQFET----- 1152
QY 218 EAPKPLALPGRGTRKRLSLTSTSPVSAKARCYPPRVVEGPHRQVLPTSPGKSWPSPA 277
DB 1152 -----CPVEVIRK---LEASTVP-----WGDYIQ-LETPA----- 1178
QY 278 RSPVPTAEKOLSSGKGVDSL-----SGSVCKHKPSTSLSPRONA----- 324
DB 1178 ---EYGRAIRSEKYQVYDILLKRPKMSVTCNAQPIRVSVM---RNFIEIADIWDMN 1231
QY 324 ---FQLRPFI-----ETRHFLYS-RGDGQERLNPFSLLS 353
DB 1232 VGSLEPFLIMLEDGDSILYDVLFTIPDIVGHEFTLSFTYELKQHNQNNLPNFFLT 1291
QY 354 NLQPNLTGARRLVEIFIGSRPRTSGPLCRHRLSRRYQWRPFLQQLLVNHAEC---QY 410
DB 1292 LISENWHSFEPIVSFNG-----FKLPKPKPPPTPLENISISTSELGNDNF 1399
QY 411 VRLLSHCFRFANQ---QVTDAL-----NTSPHMLDLRLHSSPWQ----- 451
DB 1340 SEYF-----EFKTFNKIQSQVFSLSNDSNVFVSGSGTKGTAMAEALLNH--WRQNG 1393
QY 451 --VY-----GFLRACLCKVVSASLWGRTRNRRFFKNLKKFTSLCKYKLSIQELMW 500
DB 1394 RAVYINPSGEKIDFL-----LSDMNKRFSLAGGKIINK---LGNDFSLNLKLL-- 1440

QY 501 KMKVEDCHLRSSPGKDRVPAEAHRLRERILATFLWMDTYVQ-----LLRSF 550
DB 1440 ----AKSHVLLATPVQFELLSSRRWRQKNIQSLMIYDDAHEISQGVYGVAVETLISM 1495
QY 551 FYITESTFOKNRLFFYKSV-----W-----SKLOSIGVROHLERVLRELSEQEVYR 597
DB 1496 IFIATOLEKKIREVCLSNCLANARDFGEWAGMTKSNINYFSPSERIEPLEINIQSKDVE 1555
QY 598 HHQDTWLAMPICRLRFPKPNGLRPLVNNYSYMGTRALGRKQAOHFTQRLKTL----- 652
DB 1556 HISFNF-----SMLQAFASAAAAAGNRSSSVFLPSRDCMEVASA 1597
QY 652 -----FSLNLYE-----RTKPHLMGSSVLGM--ND---IYRTWR 681
DB 1598 FMKFSKAIENDMLNVEEQIVPYIEKLTGHLRAPLKHGVILYKGMASNDERIVKRLYE 1657
QY 682 AFVLVRALDQTPRMFVKADV-----TGAYDAIPQGLVEVVANMRHSESTYCIROYA 736
DB 1658 YGAVSVLLISKDCSAFACKTDEVIILGTNLYDG-----AEHKYMPYTTINELL 1704
QY 737 VVRDSOGVHKFRROVTTLSDLQPYMGQFL-----KHLQ-----DSDASALRNSVVI 785
DB 1705 EMVGLASGNDMAGKVLILLSHNMKAYKFLIEPLTESYQVYIHDTLNNEIANSII- 1764
QY 786 EQSISMNESSSLDFLHFLHRSVVKIGRCYTOCQIGPQSS--LSTLLCSLCFGDME 843
DB 1764 -----QSKODCVDFYTSYFYRRIHVNPYSYGVDRTPSHGISVFLSNLV-ETCLNDLV 1815
QY 844 NKLFAEVQRDGLLRFVDDFLVTPHL-----DQAKTFLSTLVHGV--PEYGC----- 890
DB 1816 ESSFIED-----DTEAEVTAEVNGDDEATEIISTLSNGLASHYGVSVFFTIQSF 1866
QY 890 -----MINLQKTVVNVFVPGTGGAA-----PYQLPAH 918
DB 1867 VSSLSNTSLTKNMLVYLSLTAVEFESVPLRKGDRALLVYKSLRKLPLRPFEPH 1916
RESULT 12
ID CUL2_HUMAN STANDARD; PRT; 745 AA.
AC Q13617; O00200;
DT 01-NOV-1997 (REL. 35, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
DE CULLIN HOMOLOG 2 (CUL-2).
GN CUL2.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=KIDNEY;
RX MEDLINE; 97225922.
RA PAUSE A., LEE S., WORREL R., CHEN D.Y.T., BURGESS W.H.,
RA LINEHAN W.M., KLAUSNER R.D.;
RT "The von Hippel-Lindau tumor-suppressor gene product forms a stable
complex with human CUL-2, a member of the Cdc53 family of proteins.";
RL PROC. NATL. ACAD. SCI. U.S.A. 94:2156-2161(1997).
RN [2]
RP SEQUENCE OF 95-745 FROM N.A.
RX MEDLINE; 96279828.
RA KIPREOS E.T., LANDER L.E., WING J.P., HE W.W., HEDGECOCK E.M.;
RT "cul-1 is required for cell cycle exit in C. elegans and identifies a
novel gene family.";
RL CELL 85:829-839(1996).
CC -!- FUNCTION: FORMS A STABLE COMPLEX WITH THE VHL TUMOR SUPPRESSOR.
CC -!- SIMILARITY: BELONGS TO THE CULLIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X82939; G804981; -
KW REPEAT.
FT DOMAIN 148 272 HEAT REPEATS DOMAIN 1.
FT DOMAIN 701 898 HEAT REPEATS DOMAIN 2.
FT DOMAIN 1527 1568 HEAT REPEATS DOMAIN 3.
FT DOMAIN 18 21 POLY-GLN.
FT DOMAIN 679 682 POLY-ALA.
FT DOMAIN 1104 1108 POLY-SER.
SQ SEQUENCE 3148 AA; 348932 MW; CBB3AA6A CRC32;

Query Match 1.6%; Score 97; DB 1; Length 3148;
Best Local Similarity 19.0%; Pred. No. 1.4e+02;
Matches 244; Conservative 177; Mismatches 503; Indels 358; Gaps 64;

QY 33 PEGRRVLPQDPKPIYRTLVAAQCLVCMHMGSGPPADLSFHOYSSL--KELVARVQRLCE 90
DB 1046 PE-RRTLTGTANMVLSSA-----WF-----PDLSSAQDALLCGNLLAAVAPK-CL 1094

QY 91 RNERNVLAFCFELLNEARGGPPMATSSVRSYLPNTVETLRVSGAMLLLSRVGDDLLV 150
DB 1095 RN-----EWAGEDSSSSSTNTSGGTHKMEBPMAALSADRAVANVE 1135

QY 151 YLLAHCALYLLVPPSCAYOV---CGSPLOYCAITDIPWSVASYRTPRGVNFNLR 206
DB 1136 QLFSHLLKVLNI---CAHVLDTPGPPVKATLPSTNTPSL-----PIRRKGD----- 1184

QY 207 FLQOKSSRQAPK-----LALPSRGTKRHLSTSTSVPS-----AKKARCPVPRVEE 257
DB 1184 -KDAVDSSAPLSPKNGEANTGRTESTGSTAVHKSTTLGSFFYHLPYLYLVKATH 1242

QY 258 GPHROVLPTPSGKSWPSPARSP-EVPTAEKDLSSGKVSUL--SLSGSV---CCKHKPS- 312
DB 1243 ANFKVMDLHNSOERFGFLRAALDVLQLELATINDINCKVEILGYLASCFSREPTM 1302

QY 312 STSLSPPRQNAFQURPIETRHFLY--SRGDGQ-RLNPSFLLSNLPNI----- 360
DB 1303 ATVCVQQLKTLFGTNLASQVEGFLSGPSRSGKALRGS-----SSURPGLYHYCFMAY 1358

QY 360 -----TGARRLVEIFLGSRRPTSGPLCRTHRLSRRY--OMRPLFQQLLVNHAEC 408
DB 1359 THFTQALADASLRNMVQ-----AEHQDTSQWFDVMQKTSNQLRSNTANA 1403

QY 409 QYR-----LLRSHCR-----FRANQOVDALNTSPHLMDLRLHSSPNQVYGFRLAC 458
DB 1404 ARHRGDKNAIHNLIRLEPLVLIKALKQYTTSTVALORQVLDLLA-----QLVQ-LRVN 1456

QY 459 LCKVVSASLW-----GTRNERRFFKNLKFISLKGKYSKLQELMWMK-V 504
DB 1457 YCLDSDQVFIGFVLKOFYEYEVQFRDSEAIIPNIEFFLVLLSYERYHKSQIISIPKII 1516

QY 505 EDCHLRSSPGK---DRVPAAE---HRL-----RERILATFLWMDT 541
DB 1517 QLCOGIWSGRKAVTHAIPALQPIVHDVLVRGSKNADAGKELETOKEVYVSMLLRLVQY 1576

QY 542 YVOLLRSFFYITESTQKNRFFYKRSVMSKQSIGVROHLERLRLRELQSQEVRRHOD 601
DB 1577 H--QVLEMFVLQOCHKEN-----EDKKRLS-----ROIADVILPMIAKQOM---HLD 1621

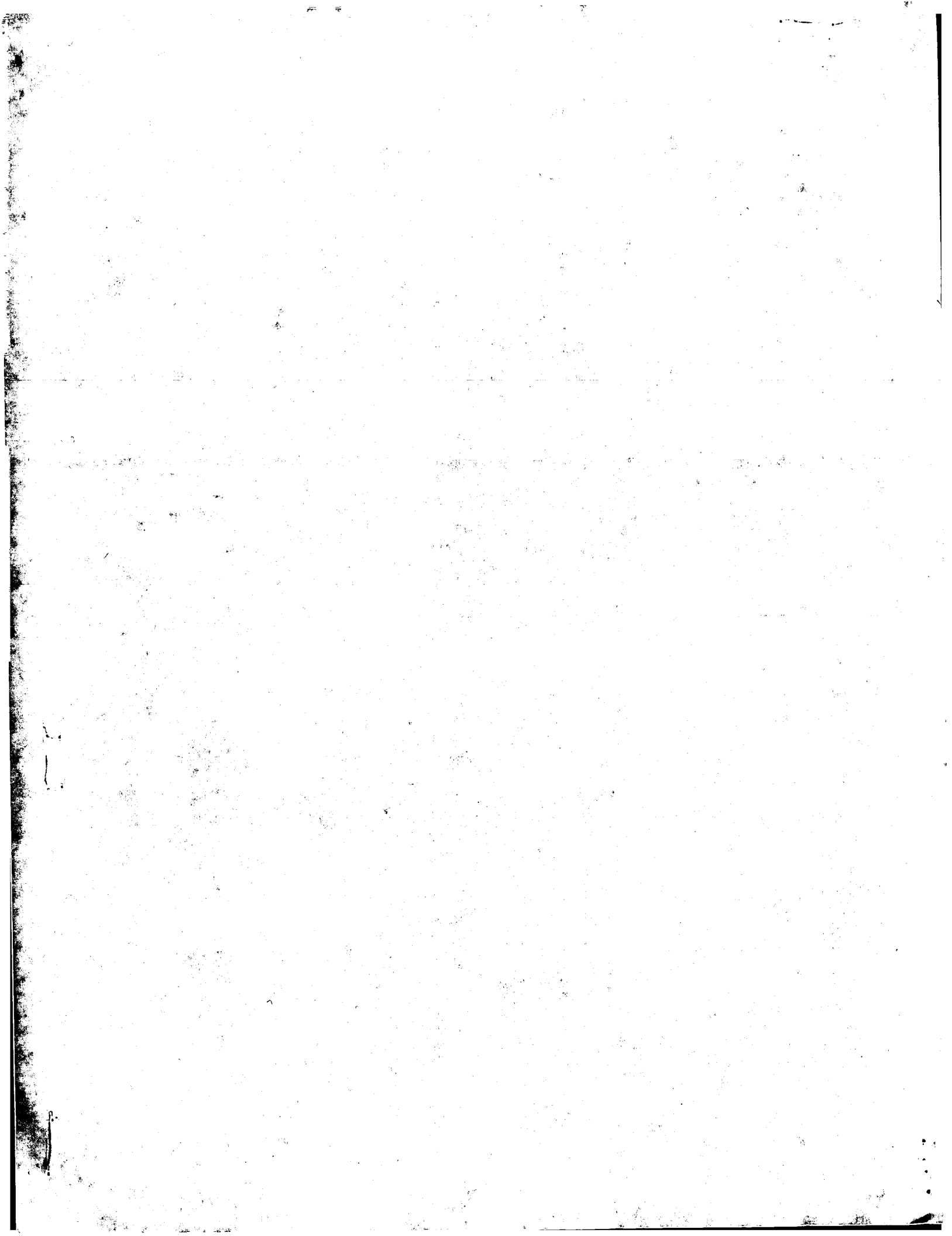
QY 602 TWLAMPICRLRF-IPKPNGLRPVNMYSMGTRALGRKQA-----QHFTORLKLTFSM 654
DB 1622 SPEALGVNLTFTETVAPSSLRPVDMLKSMFTTPVTMASVATVQLMWVSGILAVLRVLSQ 1681

QY 655 LN-----YERTKHPHLMGSSVLGNNDIYRTWRAFLVLRALDQTPRMVFKADVGA 706

DB 1682 STEDIVLSRIHELISLSPHLLSCHT-----IKRLQQPNLSPSDQPDAGDQ 1725
QY 707 YDAIPQGG-----LVEVANNIRHSESTYCIROYAVVRRDSOGVHKSPRQ 753
DB 1726 QNOEPNGEAKSLPEETARFALIOLVGLLDDISSRH-----VKVDITEQQHTFFCQ 1778
QY 754 VTTLSDLQPYMGQFLKHQDSASALRNSVIEQISNMNESS--SLFDEF---LHFLR 807
DB 1779 LGTLL-----MCLIHVFKS--GMFRITVAASRLKLGESGSHGTEFFYPLEGLNSMV 1829
QY 808 H-----SVVKIGD---RCYTQCOGIPGSSLS--TLLCSLCFGDMENK-- 846
DB 1830 HCLITHPSLVLLWCQVLLIIDYTNYSWIEVHOTPGHLSLCTKLLSPHSSGEGEKEPE 1889
QY 846 -----LFAEVORDGELLRFVDVDFLLVTPHLDQAKTFTSLVHGVPYEGCMINLQKTVNF 900
DB 1890 TRLAMINREIVRGALILFC-DYVCQNLHDSHETWL-----IVNHVRDLIDL 1936
QY 901 PVEPTLGGAAPYO---LPAHCLFPWCGLLDQTLEVFCDYSGYAQTSIKTSLTFOSVF 957
DB 1937 SHEP-----PVQDFISAVHRNSAASGLFI--QAISQRCD-----NLNSPTML 1976
QY 958 KAGTMRNKLKSLVRLKCHGLFLDLQVNSLQTVGINY-----KIFLLQAYRFA 1007
DB 1977 K--KTL--QCLEGIHLSQSGSLLMLYVDKLLSTPFRVLARMDTLACRRVEMLLAETLQ 2032
QY 1008 CVIOLPED-----QVRKNLTFFLGIITSSQASCCYAILKYKNPMTL 1049
DB 2033 SVALPLEELHRIQYLOTSLGAQRHQRFYSLLDRFRATVSDTSS-----PSTPVTSH 2085
QY 1050 KASGFPEAAHMLCYQAFLLKLAHSAVYKCLLGP---LRTAQKLLCRKLPEATWTLK 1106
DB 2086 PLDGDPPPAPELVADKEWYVALVKSOC---CLHGDVSLLETTE---LLTKLPPADLLSV 2140
QY 1107 AAD-----PALSTDFQTL 1121
DB 2141 SCKEFNLSLLCPLCSLGVQRL 2162

Search completed: August 14, 1999, 05:00:13

Job time: 537 sec



GenCore version 4.5
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 13, 1999, 15:10:16 ; Search time 18.23 Seconds
(without alignments)
3787.820 Million cell updates/sec

Title: US-09-042-460-2
Perfect score: 5901
Sequence: 1 MTRAPRCPAVRSLLRSRYRE.....TILKAAADPALSTDFQITLD 1122
Scoring table: BLOSUM62

Searched: 201082 seqs, 61543640 residues

Database : SPTREMBL_10:*
1: sp-archaea:*
2: sp-bacteria:*
3: sp-fungi:*
4: sp-human:*
5: sp-invertebrate:*
6: sp-mammal:*
7: sp-mhc:*
8: sp-organelle:*
9: sp-phage:*
10: sp-plant:*
11: sp-rodent:*
12: sp-virus:*
13: sp-unclassified:*
14: sp-unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5901	100.0	1122	11	070372
2	3505	59.4	1132	4	014746
3	3498	59.3	1132	4	014783
4	1086.5	18.4	523	4	094807
5	624	10.6	989	3	013338
6	621.5	10.5	988	3	013339
7	426.5	7.2	1132	5	076332
8	410.5	7.0	1031	5	000939
9	401.5	6.8	1117	5	077448
10	355	6.0	67	11	035432
11	341	5.8	884	3	006163
12	136.5	2.3	425	2	052331
13	133.5	2.3	561	5	045321
14	131	2.2	778	4	012815
15	129	2.2	502	5	022002
16	125	2.1	1529	10	081068
17	121	2.1	2088	10	023529
18	120.5	2.0	1646	5	019338
19	117	2.0	1569	10	022997
20	116.5	2.0	544	8	099969
21	116	2.0	945	4	094821
22	115.5	2.0	2157	3	000333
23	114.5	1.9	1159	4	014679
24	113.5	1.9	653	11	P97543
25	112	1.9	1963	6	Q28019
26	111.5	1.9	1417	12	Q57631
27	111.5	1.9	2606	12	Q36414
28	110	1.9	907	12	Q69021
29	109.5	1.9	2493	3	O60055

30	109.5	1.9	710	3	074535	074535 schizosacch
31	108.5	1.8	351	10	Q39492	Q39492 chlamydomon
32	108.5	1.8	1090	11	Q92286	Q92286 mus musculu
33	108.5	1.8	1143	12	Q10896	Q10896 hog cholera
34	107.5	1.8	601	3	Q93956	Q93956 emericeila
35	107.5	1.8	3898	12	Q09710	Q09710 pestivirus
36	107	1.8	949	3	Q14093	Q14093 schizosacch
37	107	1.8	1327	4	Q60859	Q60859 homo sapien
38	107	1.8	510	10	Q04210	Q04210 arabidopsis
39	107	1.8	1003	10	Q23218	Q23218 arabidopsis
40	107	1.8	1456	10	Q92794	Q92794 arabidopsis
41	106.5	1.8	542	2	Q68872	Q68872 myxococcus
42	106.5	1.8	1256	10	Q23528	Q23528 arabidopsis
43	106	1.8	584	2	Q08724	Q08724 frenyella d
44	106	1.8	1115	5	Q27764	Q27764 plasmodium
45	106	1.8	2528	5	Q19317	Q19317 caenorhabdi

ALIGNMENTS

RESULT 1

070372 ID 070372 PRELIMINARY; PRT; 1122 AA.
AC DT 01-AUG-1998 (T-EMBLrel. 07, Created)
DT 01-AUG-1998 (T-EMBLrel. 07, Last sequence update)
DT 01-NOV-1998 (T-EMBLrel. 08, Last annotation update)
DE TELOMERASE REVERSE TRANSCRIPTASE.
GN TERT.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia;
OC Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
RN [1]
RA GREENBERG R.A., ALLSOPP R.C., CHIN L., MORIN G.B., DEPINHO R.A.;
RL Oncogene 0:0-0(1998).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE: 98393668.
RA MARTIN-RIVERA L., HERRERA E., ALBAR J.P., BLASCO M.A.;
RT "Expression of mouse telomerase catalytic subunit in embryos and
RT adult tissues.";
DR EMBL; AF051911; AAC09323.1; -;
DR EMBL; AF073311; AAC34821.1; -;
DR MGD; MGI:1202709; TERT.
KW RNA-directed DNA polymerase.
SQ SEQUENCE 1122 AA; 127977 MW; 222075D6 CRC32;

Query Match 100.0%; Score 5901; DB 11; Length 1122;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1122; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MTRAPRCPAVRSLLRSRYREVMPLATEFVRLGPEGRRLVQGPDKIVRTLVLAOCLVCMHW 50
Db 1 MTRAPRCPAVRSLLRSRYREVMPLATEFVRLGPEGRRLVQGPDKIVRTLVLAOCLVCMHW 60
Qy 61 GSQPPPADLSFHQVSSLSKELVARVVOQLCERNRNVLAFGFELLNARGPPMAFTSSVR 120
Db 61 GSQPPPADLSFHQVSSLSKELVARVVOQLCERNRNVLAFGFELLNARGPPMAFTSSVR 120
Qy 121 SYLPTVETIETLRVSGAWMLLSRVGDDLLVLLAHACALVLLVPPSCAYOVCGSPYQICA 180
Db 121 SYLPTVETIETLRVSGAWMLLSRVGDDLLVLLAHACALVLLVPPSCAYOVCGSPYQICA 180
Qy 181 TTDIWPVSASVYRTPVGRNFTNLRFLQIKSSSRQEAAPKPLALPSRGTKRHLSITSTS 240
Db 181 TTDIWPVSASVYRTPVGRNFTNLRFLQIKSSSRQEAAPKPLALPSRGTKRHLSITSTS 240
Qy 241 VPSAKKARCYVPVREEGPHROVLPTPSCKSWVPSPARSPEVPTAEKDLSSKGKVDLSL 300
Db 241 VPSAKKARCYVPVREEGPHROVLPTPSCKSWVPSPARSPEVPTAEKDLSSKGKVDLSL 300

Mon Aug 16 10:21:27 1999

Db 241 VPSAKKACYPVPRVEBPHQVLPFTSGKSWPSPARSPEVPTAEKDLSSKGVSDLSL 300
 QY 301 SGSVCKHKPSTSLSPRONAQFAPFIETRHFLYSGDQGRNPSTLLSNLOPNT 360
 Db 301 SGSVCKHKPSTSLSPRONAQFAPFIETRHFLYSGDQGRNPSTLLSNLOPNT 360
 QY 361 GARRVEIFLGSRRPTSGPLCRTHLSRRYQWMPFIQOQLLVNHAECQVRLRSRHF 420
 Db 361 GARRVEIFLGSRRPTSGPLCRTHLSRRYQWMPFIQOQLLVNHAECQVRLRSRHF 420
 QY 421 RTANQOQVTDALNTSPHMLDLRLHSSPWQVYGFRACLCKVVSASLWGHNRERFKN 480
 Db 421 RTANQOQVTDALNTSPHMLDLRLHSSPWQVYGFRACLCKVVSASLWGHNRERFKN 480
 QY 481 LKFIISLGYKLSLOELMKWKVEDCHWLRSSPGKDRVPAEHLRERILATFLWMD 540
 Db 481 LKFIISLGYKLSLOELMKWKVEDCHWLRSSPGKDRVPAEHLRERILATFLWMD 540
 QY 541 TYVQVLLRSEFFYTESTFQKNRLFFYKRSWKSLOSIGVROHLERVRLESOEVRHQ 600
 Db 541 TYVQVLLRSEFFYTESTFQKNRLFFYKRSWKSLOSIGVROHLERVRLESOEVRHQ 600
 QY 601 DTWLAMPICRLRIPKPNGLRPTVNNYSMTGRLGRRKOAQHTQRLKTLFSLMYERT 660
 Db 601 DTWLAMPICRLRIPKPNGLRPTVNNYSMTGRLGRRKOAQHTQRLKTLFSLMYERT 660
 QY 661 KPHLMGSSVLGMDIYRTWRAFLVRALDQTPRMFYKADVTGAYDAIPGKLVVVA 720
 Db 661 KPHLMGSSVLGMDIYRTWRAFLVRALDQTPRMFYKADVTGAYDAIPGKLVVVA 720
 QY 721 NMIRHSESTYCIROYAVVRDSQGVHKSFRROVTTLSLDQPYMGQFLKHLQSDASALR 780
 Db 721 NMIRHSESTYCIROYAVVRDSQGVHKSFRROVTTLSLDQPYMGQFLKHLQSDASALR 780
 QY 781 NSVTEQISINNESSSLDFEFLHRSVVKIGDRCTYOCQIGOGSSLTLLCSLCFG 840
 Db 781 NSVTEQISINNESSSLDFEFLHRSVVKIGDRCTYOCQIGOGSSLTLLCSLCFG 840
 QY 841 DMENKLFPAEVOQDGLLLRFDVDFLLVTPHLDQAKTFLSTLVHGVPEYGCMLNOKTVNF 900
 Db 841 DMENKLFPAEVOQDGLLLRFDVDFLLVTPHLDQAKTFLSTLVHGVPEYGCMLNOKTVNF 900
 QY 901 PVEPTGLGAAPYQLPALHCLFPWCGLLDLTQTEVFCDSYGAQTSIKTSLTQSVFKAG 960
 Db 901 PVEPTGLGAAPYQLPALHCLFPWCGLLDLTQTEVFCDSYGAQTSIKTSLTQSVFKAG 960
 QY 961 KTMRNKLLSVLRKCHGLFLDLQVNSLQTVNCINIKYIFLQAYRFHACVIOQLPDRVRK 1020
 Db 961 KTMRNKLLSVLRKCHGLFLDLQVNSLQTVNCINIKYIFLQAYRFHACVIOQLPDRVRK 1020
 QY 1021 NLTFEGLIISAOASCCYATLVKVPNGMTLKASGSPPEAAHWCYQAFLLKLAHSVIYK 1080
 Db 1021 NLTFEGLIISAOASCCYATLVKVPNGMTLKASGSPPEAAHWCYQAFLLKLAHSVIYK 1080
 QY 1081 CLILGLRTAQKLLCRKLPEATMTILKAAADPALSTDFDFTILD 1122
 Db 1081 CLILGLRTAQKLLCRKLPEATMTILKAAADPALSTDFDFTILD 1122
 RESULT 2
 ID 014746 PRELIMINARY; PRT: 1132 AA.
 AC 014746;
 DT 01-JAN-1998 (TEMBLrel. 05, Created)
 DT 01-JAN-1998 (TEMBLrel. 05, Last sequence update)
 DT 01-NOV-1998 (TEMBLrel. 08, Last annotation update)
 DE TELOMERASE REVERSE TRANSCRIPTASE.
 GN HTK.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 RN Eutheria; Primates; Catarrhini; Hominidae; Homo.
 RP [1]
 SEQUENCE FROM N.A.

TISSUE-KIDNEY;
 MEDLINE; 97400623.
 RA NAKAMURA T.M., MORIN G.B., CHAPMAN K.B., WEINRICH S.L., ANDREWS W.H.,
 RA LINGNER J., HARLEY C.B., CECI T.R.;
 RT "Telomerase catalytic subunit homologs from fission yeast and
 human";
 RL Science 277:955-959(1997).
 DR EMBL; AF015950; AAC51672.1; -.
 KW RNA-directed DNA polymerase.
 SQ SEQUENCE 1132 AA; 126995 MW; 2DFBEDF3 CRC32;
 Query Match 59.4%; Score 3505; DB 4; Length 1132;
 Best Local Similarity 62.4%; Pred. No. 8.3e-298; Indels 52; Gaps 13;
 Matches 719; Conservative 122; Mismatches 260;
 QY 1 MTRAPCPAVRSLSRYREVWPLATFVRRILGPEGRRILVQGDPKIYRTLVQAQCLVCMHW 60
 Db 1 MPRAPRCRAVRSLLRSHYREVLPATFVRRILGPEGRRILVQGDPAAFRALVAQCLVCPW 60
 QY 61 GSOPPADLSFHOVSSKELVARVQVORLCERNERNVLAFFGELLNEAGGPPMAFTSSVR 120
 Db 61 DARPPAAPSFROVSCLELVARVQVORLCERGAKNVLAFFGELLNEAGGPPMAFTSSVR 120
 QY 121 SYLPNTVIELRVSGAWMLLSRVGDDLLVLLAHCALYLLVPPSCAYQVCGSPLYQICA 180
 Db 121 SYLPNTVTDALRGSGAWMLLSRVGDDLLVLLAHCALYLLVPPSCAYQVCGSPLYQICA 180
 QY 181 TTIDWPSVASYSRPTFVGRNFTNLFLQIKSSSQEAPKPLALPSRGTKRHLSTSTS 240
 Db 181 ATQARPPPHAS-GPRRLG-----CERAMNHSVREAGVPLGPAPGARRRGGASRS 231
 QY 241 VPSAKKACYPVPRVEBGP-----HRQVLPPTSGKSW-VPSPARSPVEPTAEKDLSSK 292
 Db 232 LPLKPRRGAAPPEPTPVQGSWAHPGTRGPDGFCVSPAR-----PAEATSL 286
 QY 293 GKVSLSLS-GSVCKHKPSTSLSPRONAQFAPFIETRHFLYSGDQGRNPST 350
 Db 287 GALSSTRHSHPVSGRQHAGPSTSRPPRWDTPCPVYAEKTHFLYSSGD-REQLRPSF 345
 QY 351 LLSNLOPNTGARRVELIFLGSRRPTSGPLCRTHLSRRYQWMPFIQOQLLVNHAECQY 410
 Db 346 LLSLSRPSLTGARRVETIFLGSRRPMPGTPRRLPRLPQRYQWMPFIQOQLLVNHAECQY 405
 QY 411 VRLLSHCRFTANQOQVTDAL-----NTSPHMLDLRLHSSPWQVY 452
 Db 406 GYLLKTHCPRAA---VTPAAGVCAREKPGQSVAAPEEDTDPRLVQLLRQHSSPWQVY 462
 QY 453 GFLRACLVKVSASLWGHNRERFKNLKFISLGYKLSLOELMKWKVEDCHWLR 512
 Db 463 GFVRACLRLVPPGLWGRSHNRERFKNLKFISLGYKLSLOELMKWKVEDCHWLR 522
 QY 513 SPCKDRVPAEHLRERILATFLWMDTYVQVLLRSEFFYTESTFQKNRLFFYKRSWVS 572
 Db 523 SPQVGCVPAAEHLRERILATFLWMDTYVQVLLRSEFFYTESTFQKNRLFFYKRSWVS 582
 QY 573 KLOSIGVROHLERVRLESOEVRHQDTWLAMPICRLRTHLSRRYQWMPFIQOQLLVNHA 632
 Db 583 KLOSIGVROHLERVRLESOEVRHQDTWLAMPICRLRTHLSRRYQWMPFIQOQLLVNHA 642
 QY 632 KLOSIGVROHLERVRLESOEVRHQDTWLAMPICRLRTHLSRRYQWMPFIQOQLLVNHA 692
 Db 643 RTFRREKRAERLTSRVKALFVSLNRYARRRARRRARRRARRRARRRARRRARRRARRR 702
 QY 693 TPRMYFVKADVTGAYDAIPGKLVVVAWVNNIRHSESTYCIROYAVVRDSQGVHKSFR 752
 Db 703 PPFLYFVKADVTGAYDAIPGKLVVVAWVNNIRHSESTYCIROYAVVRDSQGVHKSFR 761
 QY 753 QVTTLSLDQPYMGQFLKHLQSDASALRNSVTEQISINNESSSLDFEFLHRSVVK 812
 Db 762 HVSTLTDLPQYMQRFVAHLOET--SPLRDVAVTEQSSSLNEASSGLDFVLFPMCHAVR 819
 QY 813 IGDRCTYOCQIGOGSSLTLLCSLCFMDENKLFPAEVOQDGLLLRFDVDFLLVTPHLDQ 872

[illegible][illegible]

Mon Aug 16 10:21:27 1999

<p> RX MEDLINE: 97400623. RA NAKAMURA T.M., MORIN G.B., CHAPMAN K.B., WEINRICH S.L., ANDREWS W.H., LINGNER J., HARLEY C.B., CECCH T.R.; RT "telomerase catalytic subunit homologs from fission yeast and human." RL Science 277:955-959(1997). DR EMBL: AF015783; AAC49803.1; -. KW RNA-directed DNA polymerase. SQ SEQUENCE 989 AA; 116456 MW; FDE74202 CRC32; </p>	<p> Query Match 10.6%; Score 624; DB 3; Length 989; Best Local Similarity 23.7%; Pred. No. 8.3e-46; Matches 255; Conservative 194; Mismatches 381; Indels 244; Gaps 46; </p>
<p> QY 76 SKELVARVQPLCERN---ERNVLAFGFELLNE-ARGGPPMAFTSSVRSYLPNTVIETL 131 DB 83 SQSELIANVVKOMFESFERRNLLMKGFSMNHEDFRAMHGVQNDLVSTFPNVLISIL 142 QY 132 RVSGAWMLLSRVGDDLLVYLHAICALYLLVPPSCAYQVCGSPLYQICATTIDWPSVAS 191 DB 143 E-SKNWOLLLEIIGSDAMHYLLSKSGIFEALPNDNYLOISGIPLFK----- 188 QY 192 YRTPRVGRNFTNLRLPQOIKSSRQEPAPKPLALPSRGTKRHLSTSTSV-----PS 243 DB 188 -----NNV--FEETVSKKRKTETSIQNKARKEVSWNSISISRFSIFYRSS 234 QY 244 AKKARCPVPRVEEGPHROVLPSTGKSNVPSPARSPVPTAEKDLSSKGKVDLSLS-G 302 DB 235 YKAFK-----QDLYFNHL 247 QY 303 SVCKKHPSSTLLSPRO-----NAPQLPPTIETRHFLYSRGDQGERLNPSLLNLOPN 358 DB 248 SICDRNTVMQWIFPROGLINAFQVKQL-----HKVPLVQSQSTVFP- 293 QY 359 LTGARLVEIIFGSRPRTSGPLRTHLSRRYWMORPLFQOLLVNHAEQCVYRLRSCH 418 DB 293 -----KRLKVV-----PLIEQTAKRHRIS-----LSKVVNH-YCPVI----- 326 QY 419 RFRANQOQVTDALNTSPHMLDRLHSSPQVQVGFPLRACLCCKVVSASLWGTNRHRRF- 478 DB 326 --DTHDDE-----KILSYSLAPNQVFAELRSILRVFPKPLWG---NORIFE 367 QY 478 --FNALKKPFISLQKGLSLOELMKMKVEDCHWL-----RSSPGKDRVPAARHLRERIL 531 DB 368 IILKOLETFLLKRSYESSFSLHLYMSNIKISEIWLVLGKRSNA---KMCLSDFEKRRQIF 424 QY 532 ATFLFMDTYVQVQLLSRFFIITESTFQKNRFLFYKRSVMSKLSQISGVROHLERVL--- 589 DB 425 AEFYLYNSFIPILOQSFYIETESSDLNRNRTVFRKDIWKL---CRPFIITSMKEAF 480 QY 589 RELSQEVRHHQDTW-LAMPICRLRFPKPNGLRPVNMYSYSGMTRALGRKQAOHFT-Q 646 DB 481 EKINENNVR--MDTKTTLPPAVIRLPPKNTFRLLTNLRKRLIOMGSKMKMLYSTNQ 538 QY 647 RLKTLFSLMNYERTKHPHLMG--SSVLGMN-DIYR---TWRAFVLRVALDQTPRMVFK 700 DB 539 TLRPVASILK-----HLNEESSGIPFNLEVMTKLLTEKDLKLRHMEGR--KRYFVR 589 QY 701 ADVTGAYDAIPQGLVEVVMIRHSESTCYROYAVVRDSSQGVHKSFRQVTTLSLDL 760 DB 590 IDIKSCYDRIKQDLMFIRIVKKKLDPE--FVIRKYATHATSD-RATKNFVSEAFSTFDM 646 QY 761 QPY--MGQFLAKHODSDASALRNSVTEQISISMNESSLSLDFDLFLHFLRHSVVKIGRCY 818 DB 647 VPFEKVQVLLS-MKTS-----TLFVDFVDYWTKSSSEIFKMLKEHLSGHIVKIGNSQY 699 QY 819 TQCOGIPQSGSLTLLCSLCFGDMENKLFQVQDQD-LLLRFVDDFLVTPHLDQAKTEL 877 DB 700 LQKVGIPQSGSLSSPLCHFYMEDLIDEYLSFTKKKSVLLRVDDDFLFTVYNNKDAKREL 759 QY 878 STLHVGVPEYGCMINLQKTVVNFVPEPTGLGGAAPYQLPAHCLFPWCGLLDDTQTLVEF- 937 </p>	<p> Query Match 18.4%; Score 1086.5; DB 4; Length 523; Best Local Similarity 48.7%; Pred. No. 1.1e-86; Matches 269; Conservative 45; Mismatches 171; Indels 67; Gaps 16; </p>
<p> RN SEQUENCE FROM N.A. RP TAKAKURA M., KYO S., KANAYA T., HIRANO H., TAKEDA J., YUTSUDO M., INOUE M.; and characterization of human telomerase catalytic subunit RT (HTRT) gene promoter." RL Cancer Res. 0:0-0(1999). DR EMBL: AB016767; CAB19461.1; -. KW NON-TER 523 523 SQ SEQUENCE 523 AA; 56555 MW; OF88F1C8 CRC32; </p>	<p> Query Match 18.4%; Score 1086.5; DB 4; Length 523; Best Local Similarity 48.7%; Pred. No. 1.1e-86; Matches 269; Conservative 45; Mismatches 171; Indels 67; Gaps 16; </p>
<p> QY 1 MTRAPCPAVSLRSRYREVWPLATVRLRGEGRLVQGPDKIYRTVLAQCLVCMHW 60 DB 1 MPRAPCRVRLSLRSHYREVLPATVRLRGEGRLVQGPDPAAFRALVAOCLVCPW 60 QY 61 GSQPPPADLSFHOYSKELVARVQVORLERNVLAFFELLNEARGGPPMAFTSSVR 120 DB 61 DARPPAPSPFQVYSCLELVARVQLRCLERCAGKVNLAFFGALLDARGGPPPEAFTTSVR 120 QY 121 SYLNTVETLVRSGAWMLLSRVGDDLLVYLHAICALYLLVPPSCAYQVCGSPLYQICA 180 DB 121 SYLNTVETDARGSGAWMLLSRVGDDLLVYLHAICALYLLVPPSCAYQVCGSPLYQICA 180 QY 181 TTDWPSVASVTRPVGRNFTNLRLPQOIKSSRQEPAPKPLALPSRGTKRHLSTSTSV 240 DB 181 ATQARPPPHAS-GPERRLG-----CERANHSVREAGVPLGLPAPGARRRGSGASRS 231 QY 241 VPSAKKARCPVPRVEEGP-----HQVLPPTSGKSW-VPSPARSPVPTAEKDLSSK 292 DB 232 LPLKPRPRGAPEPPTFVGSGWAHPGRTGRGSDRGFCVVSFAR-----PAEATSL 286 QY 293 GRVSDLSLS-GVCKKHPSSTLLSPRQNAFQLRP-FIETRHFLYSRGDQGERLNPSF 350 DB 287 GALSSTRHSPVSGRQHAGPSTSRPDPWDTPCPVYAEKHFYSSGD-KEQLRPSF 345 QY 351 LLSNLQPNLTGARRLVEIFIGSRPRTSG-----PLC--RTHRLSRRYWMORPLFQOLLV 403 DB 346 LLSLRLPILTGARRLVEIFIGSGCGGLPAGCPACPSATGKCGPCFWSCLGTRTSAPT 405 QY 404 NHAECQVRLLSHCRFTANQOQVTDALNTSPHMLDRLHSS---PWQVYGFELRACLC 460 DB 406 G---CSSRTAR--CELRSPOQVPS-VPGSPRALWRPRTQTPVAV-----CSC 451 QY 461 KVVYASLWGT-----HNERRFKNLKKFISLGKYGKLSLQELMKWM 502 DB 452 SASAPLAGVRLRAGLPAPAGAPRPLGLQAQRRRLRNTKFFISLGKHLKLSLQELTWKM 511 QY 503 KVEDCHWLRSPP 514 DB 512 SVRDCAWLRSPP 523 </p>	<p> RESULT 5 OL3338 PRELIMINARY; PRT; 989 AA. AC OL3338; DT 01-JAN-1998 (Tremblrel. 05, Created) DT 01-JAN-1998 (Tremblrel. 05, Last sequence update) DT 01-NOV-1998 (Tremblrel. 08, Last annotation update) DE TELOMERASE REVERSE TRANSCRIPTASE 1. GN TET1. OS Schizosaccharomyces pombe (fission yeast). OC Eukaryota; Fungi; Ascomycota; Archiascomycetes; OC Schizosaccharomycetales; Schizosaccharomycetaceae; OC Schizosaccharomycetes. RN [1] RP SEQUENCE FROM N.A. RC STRAIN-972H-; </p>

Db 760 NLSLRGFEKHFSTLEKTVINFNENSGIINNT--FENESKRMPPFTGFSVNRSLDTLL 817
 QY 937 ---CDYSGYAQTSIK-TSLTFSQVFRAGKTRMKLLSVLRKCHGLFLDLQVNSLQTV 991
 Db 818 ACPIKIDEALFNSTSVLTCKHMGKFFV--KILRSSLASFAQ-----VFIDITHNSFN 870
 QY 992 INIYKI-----FLQAYRFHACVQLP-----FDQVRKNLTFFLGISSQASC 1035
 Db 871 CNIYRLGYSMCMRAQAYLKRMDKIFIPQRMFIDLLNVIGRKIKWKLAEILGYTSRFLS 930
 QY 1036 CYAILKYNPGMT--LKASGSPPEAAHMLCYQAFLLKLAHSHVYIKCLLGPLR 1087
 Db 931 SAEVKNLFCGLMRDGLAPSKYHP-----CFEQLIYQFOSLTDLIK-----PLR 974

RESULT 6
 O13339 ID O13339 PRELIMINARY; PRT; 988 AA.
 AC O13339;
 DT 01-JAN-1999 (TrEMBLrel. 09, Created)
 DT 01-JAN-1999 (TrEMBLrel. 09, Last sequence update)
 DE TELOMERASE REVERSE TRANSCRIPTASE 1 (EC 2.7.7.-)
 GN TRT1 OR SPC29A3.14C.
 OS Schizosaccharomyces pombe (fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Archiascomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 RN [1]
 RP Schizosaccharomycetes.
 RC SEQUENCE FROM N.A.
 RX STRAIN-972.
 RC MEDLINE: 97400623.
 RA NAKAMURA T.M., MORIN G.B., CHAPMAN K.B., WEINRICH S.L., ANDREWS W.H.,
 RA LINGNER J., HARLEY C.B., CECCH T.R.;
 RT "Telomerase catalytic subunit homologs from fission yeast and
 human";
 RL Science 277:955-959 (1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-972;
 RA LYNE M., RAJANDREAM M.A., BARRELL B.G., VOLCKAERT G.;
 RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: THIS IS THE CATALYTIC SUBUNIT AND IS DIRECTLY INVOLVED
 CC -!- SUBCELLULAR LOCATION: NUCLEAR.
 DR EMBL: AF015783; AAC49802.1; -;
 DR EMBL: AL022299; CAA18391.1; -;
 DR PFAM: PF00078; rvt; 1.
 KW Transferase; RNA-directed DNA polymerase; Telomere; Nuclear protein;
 KW DNA-binding.
 SQ SEQUENCE 988 AA; 116328 MW; 39C385A7 CRC32;

Query Match 10.5%; Score 621.5; DB 3; Length 988;
 Best Local Similarity 23.6%; Pred. No. 1.4e-45;
 Matches 253; Conservative 193; Mismatches 384; Indels 243; Gaps 45;

OY 76 SKELVARVQRLCERN---ERNVLAFGFELLNE-ARGPPMAFTSSVRSYLPNTVIETL 131
 Db 83 SQELIANVVKQMFDSFERRNLKMGFSMNHEDFRAMHVNGVQNDLVSTFPNVLISIL 142
 QY 132 RSGAWMLLSRVGDDLLVYLHAICALYLLVPPSCAYQVCGSPLYOICATTDIWPVSAS 191
 Db 143 E-SKNWQLLEITIGSDAMHYLLSKGSIFEALPNNDYLOISIPLFK-----188
 QY 192 YRPTFVGRNFTNLFLOQIKSSSRQEPAPLPSRGTCKRHLSTSTV-----PS 243
 Db 188 -----NNV--FEETVSKKRRTIETITQNKSAKEVSNWSISISRFSIFRSS 234
 QY 244 AKKARCYVPVPRVEGPHQVLPFTPSGKSWVPSPARSPEVPTAEKDLSSKGVSDLSLS-G 302
 Db 235 YKFKF-----ODLYFNLR 247

QY 303 SVCCKHPSSTSLSPRQ-----NAFQLRPFIEIHRFLYSGDQERLNPSPFLSLNLPN 358
 Db 248 SICDNTVHMVLOWIFRPFGLINAFQVKQL-----HKVPLVYSQSTVVP- 293
 QY 359 LTGARLVEIIFLGSRRPTSGPLCRTHRLSRRYQWRPFLFQQLLVNHAECQVRLRSHC 418
 Db 293 ---KRLKVV-----PLIQOTAKRLHRIS-----LSKVYNH-YCPYI----- 326
 QY 419 RFTANQOVDALNTSPPHLMDLLRLHSSPWQVYGFELRACLCKVVSASLNGRHRERF- 478
 Db 326 --DTHDDE-----KILSYSLAPNQVFAELRSILVRVFPKILWG---NORIFE 367
 QY 478 --FKNLKKFTSLGKGLSLQELMMKMKVEDCHWL-----RSSPGKDRVPAAEHRLERIL 531
 Db 368 IILKDLTFLKLSRYESFSLHYLMSNIKISEIWEVLGKRSNA---KMLSDQFEKRRQIF 424
 QY 532 ATFLFWMDTYVQLLRSPFYITESTFQKNRLLFFYKRSWMSKQSIGVROHLERVRL--- 589
 Db 425 AEFYLYNSFIPILOSPFYITESTDLNRNTVYFKDIWLL---CRPFTSMKMEAF 480
 QY 589 RELSQEVRHHODTW-LAMPICRLRFIPKPNGLRPIVNMYSMTGRALGRKQAOHFTQR 647
 Db 481 EKINENNRV--MDTQKTLPPAVIRLPPKNTFRLITNLRKRLIKMGSNKKMLVSTNQT 538
 QY 648 LKTLFSLMYERTKHPHLMG--SSVLGMN-DIYR---TWRAFVLVRVALDOTPMYFVKA 701
 Db 539 LRPVASILK-----HLINESGIPNLVYMKLTFKDKLLKRMFGR--KKYFVRI 589
 QY 702 DVTGAYDAIPQGLVEVNVANMIRHSESTYCIQYAVYVRDSQGVHKSFRROVTTLSLQ 761
 Db 590 DIKSCYDIKODLFRIVKVKKLDPD--FVIRYATIHATSD-RATKNFVSEAFSYEDMV 646
 QY 762 PY--MGQELKHLQSDASALRNSVYIQSISMNESSSLFDFELHFLRHSVYKIGRCYT 819
 Db 647 PFQVQVLLS-MKTS-----TLFVDFVDYWTYKSSSEIFKMLKEHLSGHVKNIGNSOYL 699
 QY 820 QCOGIPOGSSSLTLLCSLFCGDMENKLFQVQDQ--LLRFDVDFLLVTPHLDQAKTFLS 878
 Db 700 QKVGIPGSSILSFLCHFYMEDLIDEYLSFTKKKGVLLRVVDDFLFIVNKKDAKFLN 759
 QY 879 TLVHGVPEYCMINLQKTVNFFVEPGTILGGAAPYOLPAHCLFPACGLLLDTQTLEVF-- 937
 Db 760 LSLRGFEKHFSTLEKTVINFNENSGIINNT--FENESKRMPPFTGFSVNRSLDTLLA 817
 QY 937 ---CDYSGYAQTSIK-TSLTFSQVFRAGKTRMKLLSVLRKCHGLFLDLQVNSLQTV 992
 Db 818 ACPIKIDEALFNSTSVLTCKHMGKFFV--KILRSSLASFAQ-----VFIDITHNSFN 870
 QY 992 INIYKI-----FLQAYRFHACVQLP-----FDQVRKNLTFFLGISSQASC 1035
 Db 871 CNIYRLGYSMCMRAQAYLKRMDKIFIPQRMFIDLLNVIGRKIKWKLAEILGYTSRFLS 930
 QY 1036 CYAILKYNPGMT--LKASGSPPEAAHMLCYQAFLLKLAHSHVYIKCLLGPLR 1087
 Db 931 SAEVKNLFCGLMRDGLAPSKYHP-----CFEQLIYQFOSLTDLIK-----PLR 974

RESULT 7
 O76332 ID O76332 PRELIMINARY; PRT; 1132 AA.
 AC O76332;
 DT 01-NOV-1998 (TrEMBLrel. 08, Created)
 DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
 DE TELOMERASE REVERSE TRANSCRIPTASE.
 GN Oxytricha trifallax.
 OS Eukaryota; Alveolata; Ciliophora; hypotrichs; Stichotrichida;
 OC Oxytrichidae; Oxytricha.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE: 98337940.


```
QY 1008 CVIQLPFDORVRKNLTFFGLTISSOASCCYAILK 1041
Db 931 CAKE--YKDFKKNLMSMIDLEYSKIYSVTR 962

RESULT 9
O77448
ID O77448 PRELIMINARY; PRT; 1117 AA.
AC O77448
DT 01-NOV-1998 (Tremblrel. 08, Created)
DT 01-NOV-1998 (Tremblrel. 08, Last sequence update)
DE TELOMERASE REVERSE TRANSCRIPTASE.
GN TERT.
OS Tetrahymena thermophila.
OC Eukaryota; Alveolata; Ciliophora; Oligohymenophorea; Hymenostomatida;
OC Tetrahymenina; Tetrahymena.
RN [1]
RC STRAIN=B7;
RX MEDLINE; 98337940.
RA BRYAN T.M., SPERGER J.M., CHAPMAN K.B., CECCH T.R.;
RT "Telomerase reverse transcriptase genes identified in Tetrahymena
thermophila and Oxytricha trifallax."
PL Proc. Natl. Acad. Sci. U.S.A. 95:8479-8484(1998).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE; 98337941.
RA COLLINS K., GANDHI L.;
RT "The reverse transcriptase component of the Tetrahymena telomerase
ribonucleoprotein complex."
PL Proc. Natl. Acad. Sci. U.S.A. 95:8485-8490(1998).
DR EMBL; AF062652; AAC39135.1;
DR EMBL; AF061284; AAC39140.1;
KW RNA-directed DNA polymerase.
SQ SEQUENCE 1117 AA; 133317 MW; 1B3DF5A7 CRC32;

Query Match 6.8%; Score 401.5; DB 5; Length 1117;
Best Local Similarity 21.2%; Pred. No. 2.9e-26;
Matches 173; Conservative 147; Mismatches 358; Indels 139; Gaps 28;

QY 296 SLLSLSGSCVCKKHKSSLLSPFPQNAFQLRPFTHFYLSRGDGE----- 345
Db 190 NFNENMGKATSNNONNANLSEKKEQNOYIPEIQRSQIFCYCNHMGREGPVFKSSFFNY 249
QY 345 -RLNPSFLSNLQNLGARRLVEIIFLGSRPRT--SGPLCRTHRLSPRY--WQMRPLF- 399
Db 250 SEIKKGQFQVIOEKLOGRQFINSKIRPDHPQIIKTLLEYQ-SKNFSQGEERDLFL 308
QY 399 ---QQLLVNHAECQVRLLRSHCRF-----RTANQQVTDALNTSPPH 437
Db 309 ETEKIVQNFNINFNINLLKKECKLPENYQSLKSVQKQIVQSENKANOQSCENLFS--- 366
QY 438 LMDLLRLHSSPWQYGFRLACLCVWSASLWGTNRNRRERFKNLKFIISLGKYGKLSLOE 497
Db 366 ---LYDTEISYKQITNLRQIIQNCVNPQLLG-KNKFVLEKLYEFVQMKREENQKVL 421
QY 498 LMKMKVYEDCHLWSSPGKORVPAEH-RLRERILATFLFLMDTYVQVLLRSFFVITES 556
Db 422 YICFMDVDFVFWDLKNQKFTQRRKIYSDKRTILGLIVFIINKIVIPVLRNFYITEK 481
QY 557 TFQKNRLFFYKRSVWSKLSQISGVROHLRLRLSEQEVRRH---ODTWLAMPICRLRF 613
Db 482 HKESQIFYYRKPIWLVSKLTI-----VKLEENLEKVEKLIPEDSFQKYPQCKLRI 535
QY 614 IPRNGLRPIVNNYSYSGTRALGRKQAOHFTQRLKTLF--SMLNVERKH--PHLMGSS 669
Db 536 IPKKGSRPIMTFL-----RKQKQKNLNLQNLMDSQLVFRNLKMDLQKQIGYS 586
QY 670 VLGNNDIYRTWRAFLVRVLRALDQTPRMVFKADVTGAYDAIPGKLVEYVANNIRHS--- 727
```

```
Db 587 VFDNKOISEKFAQFIEKWKNKGR-PQLYVVTLDIKKCVDSIDQMKLL-----NFFNQSDLI 641
QY 727 ESTYCIROYAVVRDSQGVHKSFRQVTTLSDLQ-----PYM-----G 765
Db 642 QDTYFINKYLLFORNKRPLQLIQOTNNLSAMEIEEEKINKKPKFMDNINPFYFNLKER 701
QY 766 QFLKHLQSDASAL-----RNSWIEQSISSMNESSSLDFLHFLHRSVYKIG 814
Db 702 QIAYSLYDDDDQILQKFKIQSDDRPFIVINQDKPCITKDIHNLKHKISQNVISFN 761
QY 815 DRCYTQCOGIPQSSLSLTLSCFGDMENK-----LFAEVRQDG---LLRFRVDDFLVLT 867
Db 762 KVKFRKRGIPQGNISGLVCSFYFGLEEBYQFLKNAEQVNSINLLRLTDDYLFIS 821
QY 868 PHLQAKFTLSTLVHGVPEYGCINLQKTVVN--FPVEPGTLGGAAPYQIPALCLPWC 925
Db 822 DSQQNALNLVOLQNCANNNGFMENDQKITNFQFPQEDYNL---EHFKISVQNECOWIG 878
QY 926 LLLDTQILEVECDYSGYAQTSIKTSLSLTFQSVFKAGTMRNKLKSLVRLKCHGLFDL--- 983
Db 879 KSIDMNTLEI-----KSIQKQTKQEQEINQINVAISIKN-LASQLNKRLSLFLNQLID 930
QY 983 ---QVANSLOTVCINIV---KIFLLQAYRHFACVIQL 1012
Db 931 YFNPINISFEGLCRLQYHHSKATVMKFPFMTKLFQI 967

RESULT 10
O35432
ID O35432 PRELIMINARY; PRT; 67 AA.
AC O35432
DT 01-JAN-1998 (Tremblrel. 05, Created)
DT 01-JAN-1998 (Tremblrel. 05, Last sequence update)
DE TELOMERASE CATALYTIC SUBUNIT (FRAGMENT).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN [1]
RP SEQUENCE FROM N.A.
RA DRISSE R., CLEVELAND J.L.;
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF029235; AAB84200.1;
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 67 AA; 8368 MW; E2A06F2B CRC32;

Query Match 6.0%; Score 355; DB 11; Length 67;
Best Local Similarity 98.5%; Pred. No. 5.5e-24;
Matches 66; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 550 FFYITESTFQKNRLFFYKRSVWSKLSQISGVROHLRLRLSEQEVRRHQTWAMPIC 609
Db 1 FFYITESTFQKNRLFFYKRSVWSKLSQISGVROHLRLRLSEQEVRRHQTWAMPIC 60
QY 610 RLRFIPK 616
Db 61 RLRFIPK 67

RESULT 11
O06163
ID O06163 PRELIMINARY; PRT; 884 AA.
AC O06163
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DE CHROMOSOME XII COSMID 8543.
GN L8543.12.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Hemiascomycetes; Saccharomycetales;
OC Saccharomycetaceae; Saccharomycetes.
```

QY	496	QELMWKMKVEDCHWLRSPPGKDRVPAAEHLRL--ERILATFLFLWLMDTYVYVQLLRSFFYI	553
		: : : : : : : : : : : : : : :	
Db	328	DSLKLKLRLKDFWLFISD----TWFTKHENLNLQAICFISWLFROLPIKPIQITFFCY	383
		: : : : : : : : : : : : : : :	
QY	554	TE--STFOKNRLFFYKRSWVSKLOSIGVQHLRVRRLRELSQEEV--RHODTFLWAMPIC	609
		: : : : : : : : : : : : : : :	
Db	384	TEISSIV--TIVFRHDTWNKLITPFIVEFKTV---LVENNVCRRHNSYTLNSFNHS	436
		: : : : : : : : : : : : : : :	
QY	610	RLRFIPKP--NGLR----PIVNVKSYSMGTRALGRKRQAOHFTQRLKTLFSLMANYERTKHP	663
		: : : : : : : : : : : : : : :	
Db	437	KMRIIPKSKSNEERIIAIPCRGADEEEFTIYKHNKNAIQTKI-----LEYLRNKR	490
		: : : : : : : : : : : : : : :	
QY	664	HLMGSVLGMNDIYRTWRAFLRV--RALDQTPRMFYKADVGTAYDAIPQGLKLVVVAN	721
		: : : : : : : : : : : : : : :	
Db	491	TSF--TKIYSPQTADRIKEEKQRLKKFNVLPELYFMKFDVKSCYDISIPRMECMRLKD	549
		: : : : : : : : : : : : : : :	
QY	722	MIRSESTYCIROY-----AVYRDSQGVHKSRROVTTLSLDLPYMGOFLK	765
		: : : : : : : : : : : : : : :	
Db	550	ALKNENGFFVRSQYVFNTNTGVLLFNVNAGSRPKPYELIYIDNRTV-----	598
		: : : : : : : : : : : : : : :	
QY	770	HLQSDASALENSVVIQSISSMNESSSLDFDFHLFRHSVVYKIGDRCYTCQGIPOGSS	829
		: : : : : : : : : : : : : : :	
Db	598	HLSNODV-----INVMEIIF-----NTALWVEDKCYVIREDLGFOGSS	635
		: : : : : : : : : : : : : : :	
QY	830	LSTLCSLCFGDM---ENKLFAEVQRDGLLLRVRDFFLVTPHLDQAKTFLSTLVHGYPE	886
		: : : : : : : : : : : : : : :	
Db	636	LSAPIVDLVYDDLLEFYSEFKASPSQDTLLKLADDFLIISTDQOOVINIKKLAMGGFOK	695
		: : : : : : : : : : : : : : :	
QY	887	YGCMLNLQKTVVNFPEPGLTGGNAPYQLPAHGLFPCWGLLDLTQTFVFCDSYGYAOTS	946
		: : : : : : : : : : : : : : :	
Db	696	YNAAKANRDKIL-----AVSSQSDDDTVIOFCAMHIFVKEVWKHSS-----	738
		: : : : : : : : : : : : : : :	
QY	947	IKTSLTFQSVFKAGTMTNKLKSLVRLKCHGLFDLQVNSLQVTCINI-----YK	996
		: : : : : : : : : : : : : : :	
Db	738	--TMNFHRSKSSKGIFSLIALFNTRISYKTTIDTLNSTNLTMQLDHRVKNISECYK	795
		: : : : : : : : : : : : : : :	
QY	997	-----IFLLQAYRFHACVQLQFPDQRY	1018
		: : : : : : : : : : : : : : :	
Db	796	SAFKDLISNVTQNMQFHS-----FLORI	818
		: : : : : : : : : : : : : : :	
RESULT	12		
O52231		PRELIMINARY;	PRT; 425 AA.
ID	O52231		
AC	O52231;		
DT	01-JUN-1998 (TREMBLrel. 06, Created)		
DT	01-JUN-1998 (TREMBLrel. 06, Last sequence update)		
DT	01-MAY-1999 (TREMBLrel. 10, Last annotation update)		
DE	MARURASE-RELATED PROTEIN.		
OS	Streptococcus pneumoniae.		
OC	Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;		
OC	Streptococcus.		
EN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN-NCIC11906;		
RX	MEDLINE; 98125733.		
RA	COFFEY T.J., ENRIGHT M.C., DANIELS M., MORONA J.K., MORONA R.,		
RA	HRYNIEWICZ W., PATON J.C., SPRATT B.G.;		
RT	"Recombinational exchanges at the capsular polysaccharide		
RT	biogenetic locus lead to frequent serotype changes among natural		
RT	isolates of Streptococcus pneumoniae."		
RL	MOI. Microbiol. 27:73-83(1998).		
DR	EMBL; AF030367; AAC38715.1; -		
DR	PFAM; PF00078; rvt. 1.		
SO	SEQUENCE 425 AA; 49537 MW; 5AD1F39A CRC32;		

```

Query Match      2.3%; Score 136.5; DB 2; Length 425;
Best Local Similarity 19.8%; Pred. No. 0.0011;
Matches 74; Conservative 48; Mismatches 106; Indels 149; Gaps 13;
QY 539 MDTYVVQLLRSFFVITESTFOKNRFLFPRKSVMSKLQ----SIGVRQ-HLEVRVRLRSQ 593
      :|:|:| :|:|:| :|:|:| :|:|:| :|:|:| :|:|:| :|:|:| :|:|:| :|:|:|

```

```

RN      SEQUENCE FROM N.A.
RP      STRAIN=S288C (AB972);
RC      JOHNSON M., HILLIER L., RILES L., ALBERMANN K., ANDRE B., ANSORGE W.,
RX      MEDLUNG M., HILLIER L., DUBOIS E., DUSTERHOFT A.,
RA      BENES V., BRUCKNER M., DELIUS H., FLOTH M., GOFFEAU A., HEBLING U., HEUMANN K.,
RA      ENTAN K.D., FLOETH M., GILBERT H., HILGER F., KLEINE K., KOTTER P.,
RA      HEUGS-NEITZEL D., HILBERT H., MEWES H.W., MIOGSA T., MOSTL D.,
RA      LOUIS E.J., MESSENGUY F., NENTWICH U., OBERMAIER B., PIRAVANDI E., POHL T.M.,
RA      MULLER-AUER S., PORNELLE B., RECHMANN S., RIEGER M., RINKE M., ROSE M.,
RA      PORTETELLE D., SCHERENS B., SCHOLLER P., SCHWAGER C., SCHWARZ S.,
RA      SCHARFE M., SCHERENS B., URRUTARAZU L.A., VANDENBOL M., VERHASSELT P.,
RA      VIERENDEELS F., VOET M., VOLCKAERT G., VOSS H., WAMBUIT R., WEDLER E.,
RA      WEDLER H., ZIMMERMANN F.K., ZOLLNER A., HANI J.J., HOEHEISEL J.D.;
RT      *The nucleotide sequence of Saccharomyces cerevisiae chromosome
RT      XII.*;
RT      Nature 387:0-0(0).
RN      [2]
RN      SEQUENCE FROM N.A.
RC      STRAIN=S288C (AB972);
RA      DU Z.;
RA      Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.
RN      [3]
RN      SEQUENCE FROM N.A.
RP      STRAIN=S288C (AB972);
RC      WATERSTON R.;
RA      Submitted (FEB-1995) to the EMBL/GenBank/DBJ databases.
RN      [4]
RN      SEQUENCE FROM N.A.
RP      STRAIN=S288C (AB972);
RC      CHERRY J.M.;
RA      Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
DR      EMBL; U20618; AAB64520.1; -
SQ      SEQUENCE 884 AA; 102662 MW; 1A94320F CRC32;

Query Match      5.8%; Score 341; DB 3; Length 884;
Best Local Similarity 19.8%; Pred. No. 4.1e-25;
Matches 208; Conservative 148; Mismatches 355; Indels 338; Gaps

QY 33 PEGRRLVQPGDPKTYRTLVLAQCLVCHWGSQPPADLSFHQVSSLLKELVARVQRLCERN 92
DB 45 PNSRKIALP-----CL-----PGDSLKHAV-----IDHCIIYLLTGEL 77
QY 93 ERNVLAFCGELL-NEARGGPPMAFTSSVRSYLPNTVIETLRVSGAWMLLSRVGDDLLVY 151
DB 78 YNNVLTFGKIARNEDVNSLFCHSANV-----VTLKGAANKMFHSLVGTYAFVD 129
QY 152 LLAHCALYLLVPPSCAYQVCGSPYQI-----CATTDIWP-----SVSASVTRTPVGR 200
DB 130 LLIN--YTVI-----QFNQOFFTQVGNRCNEPHLPKRWQVRSSSSATAAQIKLTE 180
QY 201 NFTNLRFLOQIKSSSRQEAAPLALPSRGTKRHLISLTSTVSPSAKKARCYPPRVVEGPH 260
DB 181 PVTKQFOLHK-----LNINSSSF-----PPY 201
QY 261 RQVLTPSGKSWVPSPARSPEVPTAEXDLSKGVDSLSGVCCKHKXPSSTLSLSPR 320
DB 202 SKILPSS-----SSIKLTDL-----REAFPTNLVKIP- 231
QY 321 ONAFQRLPIETHFELYSRGDOERLNPFLLSNLQPLNLTGARRLVEIIFLGSRPRTSGP 380
DB 231 -----QRL-----KVRINLT----- 241
QY 381 LCRTHLRSRYWQMRPLFOQLLVNHAEOCYVRLRLSRHCRTFANQOVTALNTSP---H 437
DB 241 -----LQKLLKRHLKNVYVILNSIC-----PPEGT 267
QY 438 LMDLLRL--HSSPWQYVGLRACLCKVVSASILWGTGRTRNERRFFKNLKFISLKGKLSL 495
DB 268 VLDLSHLSRQSPKRVLKPTIIVLOKPLQPOEMFGSKNKGKIKNLNLLSLPLNGLYLFP 327

```

Query Match	5.8%;	Score 341;	DB 3;	Length 884;
Best Local Similarity	19.8%;	Pred. NO. 4.1e-21;		
Best Local Similarity	19.8%;	Pred. NO. 4.1e-21;		
	149;	Mismatches 355;	Indels 338;	Gaps 40;

```
Db 38 MDNY-----LRQNWRLTKELIKQK---YKQPVLVKEIPKPDGGRQIGIFVMDRMIOQ 90
QY 594 EYVRHQDTWAMPICRLURFIPKPNGLRPIVNMYSMTGRALGRKQAOHQFTQRLKTLFS 653
Db 91 AIVQ-----VMSPICEPHFSDTSYGRFP--NRSC-----KAIMK 123
QY 654 MLNVERTKPHLMGSSVLGMNDIVRTWRAFLVRLALDQTPRMVYKADVTGAYDAIPQG 713
Db 134 LLEY-----LNDGYE-----WIVDIDLEKFFDTPQD 150
QY 714 KLVEVWAMIRHSSTYCIROYAVVRDSDQGVHKSFRRQVTTLSDLQPYMGQLKLQD 773
Db 151 RLMSLVNHIIDGDGTESLIRKYL-----174
QY 774 SDASALRNSVVEIOSISNNESSSLDFFLHFLRHSVVKIGDRCYTCOCGIPQGSLSLTL 833
Db 174 -----HSGVIINGORYKTLVGTPOGGNLSPL 199
QY 834 LCSLCFGDMENKLFAYVQRDGL--LLREVDVDFLLVTPHLDQAKTFLSLVHGVPF--YGCMI 891
Db 200 LSNI-----MLNLDKEKRLRVRVADDCVTVGSEAAKRVMSVSRFIEKRGLKV 255
QY 892 NLOKTVVNFVPGTGLG 908
Db 256 NMTKTKITRPRELKYLG 272
RESULT 13
O45321
ID O45321 PRELIMINARY; PRT; 561 AA.
AC O45321;
DT 01-JUN-1998 (Tremblrel. 06, Created)
DT 01-JUN-1998 (Tremblrel. 06, Last sequence update)
DT 01-JAN-1999 (Tremblrel. 09, Last annotation update)
DE DY3.4 PROTEIN.
GN DY3.4.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Secernentes; Rhabditia; Rhabditida;
OC Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.
RN [1]
RP SEQUENCE FROM N.A.
RA LENNARD N.;
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE; 94150718.
RA WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M., COULSON A.,
RA BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., FULTON L.,
RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., JIER M., JOHNSTON L.,
RA GARDNER A., GREEN P., HAWKINS T., HILLIER L., LAISTER N., LATREILLE P.,
RA JONES M., KERSHAW J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,
RA LIGHTNING J., PERCY C., RIEKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,
RA PARSONS J., SMITH A., SONNHAMMER E., STADEN R., SULTON J.,
RA SMALDON N., THERRRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,
RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans.";
RL Nature 368:32-38(1994).
DR EMBL; 296047; CAB09413.1; -.
SQ SEQUENCE 561 AA; 65807 MW; 1C710A97 CRC32;
```

Query Match 2.3%; Score 133.5; DB 5; Length 561;
Best Local Similarity 19.7%; Pred. No. 0.003;
Matches 66; Conservative 65; Mismatches 149; Indels 55; Gaps 12;

```
QY 617 PNGLRPIV---NMSYS-MGTRALGRKQAOHQFTQRL--KTLFSMLNY--ERT-KHPLMG 667
Db 249 PHFIRPNVATFKLSUSKQKLPFLKRAIDKKTETMQKRLNSMLSWCLERSGVYRHTIR 308
QY 668 SSVLGMNDIYRTWRAFLVRLALDQTPRMVYKADVTGAYDAIPQGLVEVWAMIRHSE 727
```

```
Db 309 DSCCKVYSDF-----LKNQSNKIIGYADVSKCFSTVNHVLIISIDRLFSQBH 358
QY 728 STYCIROYAVVRDSDQGVHKSFRRQVTTLSDLQPYMGQLKLQDSDASALRNSVWIEQ 787
Db 359 DIYV-----CGKRNHGGFHKLIFCSAGTELNAHEALR--RKMKLGVFNEVCYR 408
QY 788 STSMNESSSLDFFLHFLRHSVVKIGDRCYTCOCGIPQGSLSLTLCSLCFGDMENKLF 847
Db 409 EMS-----SSTILSVIRTTLSYTYKRGPTSWRITKGVPOGHPISSNLAHMYLNNEEQYK 465
QY 848 AEVQRDGLL--RFYDDVDFLLVTPHLDQAKTFLSLVHGVPYEGCMINLQKTVNFPVPG 905
Db 466 SNEKEDSRIVFCRYEDDFITFENSLFEKMMKPLSTGNHTFLTANPKK-----516
QY 906 TLGGAAPQLPAHC-----LFPWCGLLLOTQTLEVF 936
Db 516 -----FKKSRGCGASQVQWCGVKLDFQSGNCF 543
RESULT 14
Q12815
ID Q12815 PRELIMINARY; PRT; 778 AA.
AC Q12815;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-NOV-1998 (Tremblrel. 08, Last annotation update)
DE TASTIN.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 95278733.
RA FUJUDA M.N., SATO T., NAKAYAMA J., KLIER G., MIKAMI M., AOKI D.,
RA NOZAWA S.;
RT "Trophinin and tastin, a novel cell adhesion molecule complex with
RT potential involvement in embryo implantation.";
RL Genes Dev. 9:1199-1210(1995).
DR EMBL; U04810; AAA79333.1; -.
KW Cell adhesion.
SQ SEQUENCE 778 AA; 83758 MW; 4FE9DC3A CRC32;
```

Query Match 2.2%; Score 131; DB 4; Length 778;
Best Local Similarity 22.4%; Pred. No. 0.008;
Matches 107; Conservative 44; Mismatches 167; Indels 160; Gaps 26;

```
QY 2 TRAPRCPAVRSLLRSRYRVMPLATFVRRLGPGRLVQPGDPKIIYRTLVA-----QC 54
Db 358 TPWPSTPRVQ-----AQLRGVSPQS---CSEDPALPWEQVAVRLPDQESC 401
QY 55 LYCMHMGSQPP---PADLSFHVSSLKELVAR---VVORLCERNVNLAFGFEELNEAR 108
Db 402 IRSLEGGSGKPPVATPSGPHSNRTPSLOEVKIQIGILQQLLROEVLGGQCVPLN---459
QY 109 GGP--PWAFSTSVRSVRLPNTVIETLRVSGAMWMLLSRVGDDLLVYLHAHALYLLVPPSC 166
Db 459 GGSLLDMVQLPLLTETISRLTNAHNSGT-----SHLPGLLKHSLG---PKPC 504
QY 167 AYOVCGSPLYQICATTDIWP-----SVSASVYRTPRPVGRNFTN-L 205
Db 505 LPECEGP--QPCPPAEPGPPEAFRCSEPEIPEPSLQEOLEVPYPPAEP--RPLESCC 560
QY 206 RFLQOIKSSRQF-----APKPL-----ALPSRGTKRHSLTSTSVPSAKK 246
Db 561 RSEPEIPESSRQEOLEVPPEPCPPAEPRLPESYCRLEPEIPESRQEL-----EYPE---613
QY 247 ARCYPVPRVEEGPHROVLTPTSGKSWVSPARSPVEPTAEKDLSSKGVSDLSLGSV--305
Db 613 -----PCPPAEPGP---LQPSIQSGSGPPGC-----PRVELGASEPCTLEHRSLESSLPP 660
QY 305 CCKH-KPSSTSL-----SPRONAFQLRPFIEIRHFLYSRGDCQERLNPFSLLSN 354
```

Mon Aug 16 10:21:27 1999

375 A---IYGRFTDKRPHKVLTLHETTNEAAQCLLVPSLL--TRRDE--LFPLARQIVK 427

Db 661 CCSQAPATTSIFSSQHPCLASPPICSLQSRP-----PAGQAG-----LSN 703
QY* 355 LQNLGTGARRLV-----EIIFLGSRPRTSGP--LCRTHRLSRWYQ 393
Db 704 LAPRTIALRESLKSCLTAIHCHEARLDDCAFYTSRASPSGPTRVCTNPVATLEWQ 761

QY 487 LGKYGKLSLQELMWKMKVEDCHWLRSPPGKDRVP 520
Db 428 DAGYNYAKSKR--PCDPADLHSPISSPGNSPPP 459

Search completed: August 13, 1999, 17:33:42
Job time: 8606 sec

RESULT 15
Q22002 ID Q22002 PRELIMINARY; PRT; 502 AA.
AC Q22002;
DT 01-NOV-1996 (TREMELrel. 01, Created)
DT 01-NOV-1996 (TREMELrel. 01, Last sequence update)
DT 01-JAN-1999 (TREMELrel. 09, Last annotation update)
DE RI66.1 PROTEIN.
GN RI66.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;
OC Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.
RN [1]
RP SEQUENCE FROM N.A.
RA MATTHEWS P.;
RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE: 94150718.
RA WILSON R., AINSKOUGH R., ANDERSON K., BAYNES C., BERKS M.,
RA BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., COULSON A.,
RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,
RA GARDNER A., GREEN P., HAWKINS T., HILLIER N., JIER M., JOHNSTON L.,
RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,
RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,
RA PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,
RA SMALDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,
RA THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,
RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RL Nature 368:32-38(1994).
DR EMBL: Z50795; CAA90662.1; .
SQ SEQUENCE 502 AA; 53782 MW; D037882A CRC32;

Query Match 2.28; Score 129; DB 5; Length 502;
Best Local Similarity 22.0%; Pred. No. 0.0063;
Matches 100; Conservative 57; Mismatches 159; Indels 138; Gaps 22;
QY 141 LSRVGDLLVLLARCALYLLVPPSCAYQVCGSPLYQICATTD---IWPSVSASYPTR 196
Db 70 LAEYSQDOTAFNLA--ALQOIGPP-----PPLNYTPAGTDPALLPGIAAATSP-- 118
QY 197 PVGRNFTNRLFLQIKSSSRQAPKPLALPSRGTKRHLSTSTSVPSAKKARCPVPRVE 256
Db 118 ---KFPSLRFLSQLSSVAEKVSTPSEADTSSSPSLNLTSSNSVPLAFKFTPLLE 173
QY 257 E-----GPHRQVLTPTSGKSWPS-----PARSPEVPTAEKDLSSK-----G 293
Db 174 SLASDQSSSTSVRSVLPSSTNSHPELPAGILPATTVNSAAVPPSSRATANVFSG 233
QY 294 KYSDLSLGSV--CCKH---KPSSTSLSP-----RQNAFQLRPFETRHFLYS 338
Db 234 NSIGLNFSGAASVTRHLVVPSTSTQOPSTSGFRSSSITGOEKGSSSPFLGVG--YS 290
QY 339 RGDQERLNPFLSLNQLNLTGARRLVEIIFLGSPPRTSGPLCRTHRLSRRYWQMRPLF 398
Db 291 QPYGNDVFS---LGDFDN-----NPSLTENPTLTAQISRL----- 325
QY 399 QQLLVNHAEOYV-----RLRSHCRFTANQOVTDAINTSP--PHLMDLLRLHS 446
Db 325 -----AECALAAKNLPLPLRLVQNKRR---VSKEVIELLKCSPATPSMIHAFKYS 374
QY 447 SPWQVYGLF-----RACLCKVVSASLWCTRHNERREFKNLKKFIS 486